

Job time : 115 53 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:42:54 ; Search time 44.7332 Seconds
(without alignments)
640.518 Million cell updates/sec

Title: US-09-662-293-15

Perfect score: 3107
Sequence: 1 MKTIVAILSMACIGLNMNAS.....IMCPPTGTCOEKLTICGE 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/aaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/aaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/aaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/aaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/aaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------|-------------------|
| 1 | 3107 | 100.0 | 555 | US-09-292-225-15 | Sequence 15, Appl |
| 2 | 3107 | 100.0 | 555 | US-09-292-225-18 | Sequence 18, Appl |
| 3 | 3014 | 97.0 | 536 | US-09-292-225-21 | Sequence 21, Appl |
| 4 | 2542 | 81.8 | 509 | US-09-292-225-35 | Sequence 35, Appl |
| 5 | 2542 | 81.8 | 509 | US-09-292-225-38 | Sequence 38, Appl |
| 6 | 2475 | 79.7 | 490 | US-09-292-225-41 | Sequence 41, Appl |
| 7 | 838.5 | 27.0 | 554 | US-08-524-051-2 | Sequence 2, Appl |
| 8 | 838.5 | 27.0 | 554 | US-09-052-778-16 | Sequence 16, Appl |
| 9 | 815.5 | 26.2 | 583 | US-09-545-814-2 | Sequence 2, Appl |
| 10 | 815.5 | 26.2 | 583 | US-09-545-814-5 | Sequence 5, Appl |
| 11 | 812.5 | 26.2 | 635 | US-09-545-814-32 | Sequence 32, Appl |
| 12 | 811.5 | 26.1 | 559 | US-09-545-814-14 | Sequence 14, Appl |
| 13 | 738 | 23.8 | 466 | US-08-486-839-4 | Sequence 4, Appl |
| 14 | 738 | 23.8 | 466 | US-09-151-011-4 | Sequence 4, Appl |
| 15 | 738 | 23.8 | 466 | US-09-039-198A-2 | Sequence 2, Appl |
| 16 | 738 | 23.8 | 466 | US-09-343-623-4 | Sequence 4, Appl |
| 17 | 738 | 23.8 | 466 | US-08-877-599-2 | Sequence 2, Appl |
| 18 | 738 | 23.8 | 466 | US-09-267-574-2 | Sequence 2, Appl |
| 19 | 732.5 | 23.6 | 387 | US-08-486-839-6 | Sequence 6, Appl |
| 20 | 732.5 | 23.6 | 387 | US-09-151-011-6 | Sequence 6, Appl |
| 21 | 732.5 | 23.6 | 387 | US-09-039-198A-4 | Sequence 4, Appl |
| 22 | 732 | 23.6 | 466 | US-09-343-623-6 | Sequence 6, Appl |
| 23 | 732 | 23.6 | 466 | US-08-877-599-4 | Sequence 4, Appl |
| 24 | 732 | 23.6 | 466 | US-09-267-574-4 | Sequence 4, Appl |
| 25 | 721.5 | 23.2 | 373 | US-09-039-198A-14 | Sequence 14, Appl |
| 26 | 721.5 | 23.2 | 373 | US-09-039-198A-15 | Sequence 15, Appl |
| 27 | 721.5 | 23.2 | 373 | US-08-877-599-14 | Sequence 14, Appl |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 721.5 | 23.2 | 373 | 4 | US-08-877-599-15 | Sequence 15, Appl |
| 29 | 721.5 | 23.2 | 373 | 4 | US-09-267-574-14 | Sequence 14, Appl |
| 30 | 721.5 | 23.2 | 373 | 4 | US-09-267-574-15 | Sequence 15, Appl |
| 31 | 672.5 | 21.6 | 489 | 4 | US-09-545-814-29 | Sequence 29, Appl |
| 32 | 650 | 20.9 | 383 | 4 | US-09-459-749D-17 | Sequence 17, Appl |
| 33 | 630 | 20.3 | 385 | 2 | US-08-694-915-2 | Sequence 2, Appl |
| 34 | 630 | 20.3 | 416 | 2 | US-08-694-915-4 | Sequence 4, Appl |
| 35 | 630 | 20.3 | 423 | 4 | US-08-850-348A-2 | Sequence 2, Appl |
| 36 | 420 | 13.5 | 442 | 3 | US-09-052-778-2 | Sequence 2, Appl |
| 37 | 419.5 | 13.5 | 700 | 4 | US-09-408-647A-2 | Sequence 2, Appl |
| 38 | 417 | 13.4 | 440 | 3 | US-09-052-778-13 | Sequence 13, Appl |
| 39 | 412 | 13.3 | 423 | 1 | US-07-939-501A-10 | Sequence 10, Appl |
| 40 | 412 | 13.3 | 423 | 1 | US-07-939-501A-12 | Sequence 12, Appl |
| 41 | 411 | 13.2 | 389 | 1 | US-07-939-501A-1 | Sequence 1, Appl |
| 42 | 411 | 13.2 | 389 | 3 | US-08-448-398-7 | Sequence 7, Appl |
| 43 | 403.5 | 13.0 | 561 | 1 | US-08-358-901-2 | Sequence 2, Appl |
| 44 | 403.5 | 13.0 | 561 | 1 | US-08-566-347-2 | Sequence 2, Appl |
| 45 | 403.5 | 13.0 | 561 | 1 | US-08-693-835-2 | Sequence 2, Appl |

ALIGNMENTS

RESULT 1
US-09-292-225-15
; Sequence 15, Application US/09292225
; Patent No. 6455866
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-15

| | | | | | | | |
|-----------------------|---------|---|---------------------------|------------|----|--------|------|
| Query Match | 100.0%; | Score | 3107; | DB | 4; | Length | 555; |
| Best Local Similarity | 100.0%; | Pred. NO. | 6.5e-241; | | | | |
| Matches | 555; | Conservative | 0; | Mismatches | 0; | Indels | 0; |
| Gaps | 0; | | | | | | |
| Oy | 1 | MKTIVAILSMACIGLNMNASIKKDHNDYKKNPVI | CVGCTWSYVHKVDVPTTIEDIDPFK | 60 | | | |
| Db | 1 | MKTIVAILSMACIGLNMNASIKKDHNDYKKNPVI | CVGCTWSYVHKVDVPTTIEDIDPFK | 60 | | | |
| Oy | 61 | CTHLMYGFAXIDEXKTYIQVDFPYQDDNHNSWEKRGYERFNNLR | LKNPELTMTISLGGWY | 120 | | | |
| Db | 61 | CTHLMYGFAXIDEXKTYIQVDFPYQDDNHNSWEKRGYERFNNLR | LKNPELTMTISLGGWY | 120 | | | |
| Oy | 121 | EGSKYSDMAANPTYRQFTQSVLDFLOEYKFGDLDDWEYPSGRLGNPKIDKQNYALV | 180 | | | | |
| Db | 121 | EGSKYSDMAANPTYRQFTQSVLDFLOEYKFGDLDDWEYPSGRLGNPKIDKQNYALV | 180 | | | | |
| Oy | 181 | RELKDAPEPHGYLLTAASVSGKIDRAYDIKEINKLFDWMNTYDYHGGWENFYGHNA | 240 | | | | |
| Db | 181 | RELKDAPEPHGYLLTAASVSGKIDRAYDIKEINKLFDWMNTYDYHGGWENFYGHNA | 240 | | | | |
| Oy | 241 | PLYKRPTDELHTYFNVNVTMTYLLNNGATRDKLVMGVFPYGRAWSIEDRSKLKLGDA | 300 | | | | |

Db 241 PLYKRPDETDLHTYFNNTYTHYLLNNGATRDKLVGVFPYGRWSIEDRSKLGDP 300
Qy 301 KMSPPGFIISGEGVLSYIELCOLFQKEWHIQYDEYNAPYGVNDKIWGYDDLASISC 360
Db 301 KMSPPGFIISGEGVLSYIELCOLFQKEWHIQYDEYNAPYGVNDKIWGYDDLASISC 360
Qy 361 KLAFLKELGVSGVMVWSLENDDFKGHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
Db 361 KLAFLKELGVSGVMVWSLENDDFKGHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
Qy 421 TPPTTPT 480
Db 421 TPPTTPT 480
Qy 481 TPSPTTTEHTSETPKYTYVDGHLIKYKKGDIHPHTNIHKYLVCFVNGVWVHMP 540
Db 481 TPSPTTTEHTSETPKYTYVDGHLIKYKKGDIHPHTNIHKYLVCFVNGVWVHMP 540
Qy 541 PGTIWCQEKLTICGE 555
Db 541 PGTIWCQEKLTICGE 555

RESULT 2

US-09-292-225-18
; Sequence 18, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-18

Query Match 100.0%; Score 3107; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKTYAIIISIMACIGLNNASIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
Db 1 MKTYAIIISIMACIGLNNASIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
Qy 61 CTHLMYGFADIKYKTIQVDFPDYODDNHNSWEKRGYERFNNLKNPELTMTISLGGWY 120
Db 61 CTHLMYGFADIKYKTIQVDFPDYODDNHNSWEKRGYERFNNLKNPELTMTISLGGWY 120
Qy 121 EGSEKYSMAANPTVROQFIQSVLDFLQYKFDGLDLDWEYPSGLNPKIDKONYLALV 180
Db 121 EGSEKYSMAANPTVROQFIQSVLDFLQYKFDGLDLDWEYPSGLNPKIDKONYLALV 180
Qy 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNA 240
Db 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNA 240

Qy 241 PLYKRPDETDLHTYFNNTYTHYLLNNGATRDKLVGVFPYGRWSIEDRSKLGDP 300
Db 241 PLYKRPDETDLHTYFNNTYTHYLLNNGATRDKLVGVFPYGRWSIEDRSKLGDP 300
Qy 301 KMSPPGFIISGEGVLSYIELCOLFQKEWHIQYDEYNAPYGVNDKIWGYDDLASISC 360
Db 301 KMSPPGFIISGEGVLSYIELCOLFQKEWHIQYDEYNAPYGVNDKIWGYDDLASISC 360
Qy 361 KLAFLKELGVSGVMVWSLENDDFKGHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
Db 361 KLAFLKELGVSGVMVWSLENDDFKGHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
Qy 421 TPPTTPT 480
Db 421 TPPTTPT 480
Qy 481 TPSPTTTEHTSETPKYTYVDGHLIKYKKGDIHPHTNIHKYLVCFVNGVWVHMP 540
Db 481 TPSPTTTEHTSETPKYTYVDGHLIKYKKGDIHPHTNIHKYLVCFVNGVWVHMP 540
Qy 541 PGTIWCQEKLTICGE 555
Db 541 PGTIWCQEKLTICGE 555

RESULT 3

US-09-292-225-21
; Sequence 21, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-21

Query Match 97.0%; Score 3014; DB 4; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.8e-233;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 SIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 79
Db 1 SIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 60
Qy 80 VFDPYODDNHNSWEKRGYERFNNLKNPELTMTISLGGWEGSEKYSMDAANPTVROOF 139
Db 61 VFDPYODDNHNSWEKRGYERFNNLKNPELTMTISLGGWEGSEKYSMDAANPTVROOF 120
Qy 140 IQSVLDFLQYKFDGLDLDWEYPSGLNPKIDKONYLALVRELKDAPEPHGYLLTAAVS 199
Db 121 IQSVLDFLQYKFDGLDLDWEYPSGLNPKIDKONYLALVRELKDAPEPHGYLLTAAVS 180
Qy 200 PGKDKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDLHTYFN 259
Db 181 PGKDKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDLHTYFN 240

QY 260 YTHYVYLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPKMSPPGISGEGVLSVI 319
DB 241 YTHYVYLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPKMSPPGISGEGVLSVI 300
QY 320 ELQCLFQKEBWHIOYDEYVYNAFYNDKIWVGDDLASISCKLAFKELGVSGVWWSLE 379
DB 301 ELQCLFQKEBWHIOYDEYVYNAFYNDKIWVGDDLASISCKLAFKELGVSGVWWSLE 360
QY 380 NDDPKGHCGRPNLLKVNHWINGDEKNSPECILGSPSTTPTPTPTPTPTPTPTPTPT 439
DB 361 NDDPKGHCGRPNLLKVNHWINGDEKNSPECILGSPSTTPTPTPTPTPTPTPTPTPT 420
QY 440 TPTTPT 499
DB 421 TPTTPT 480
QY 500 VDGLHLKCYKEGDIPIHPTNIHKYLVEFVNGVWVHIMPCCGTIWCQEKLTICGE 555
DB 481 VDGLHLKCYKEGDIPIHPTNIHKYLVEFVNGVWVHIMPCCGTIWCQEKLTICGE 536

RESULT 4

US-09-292-225-35
; Sequence 35, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-35

Query Match 81.8%; Score 2542; DB 4; Length 509;
Best Local Similarity 81.7%; Pred. No. 1.1e-195;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

QY 1 MKTIYALISIMACIGLMAASIKRDHNDYSKNPMRIVCYVGTWSVYHKVDPYTTIEDIDPFK 60
DB 1 MKTTFALFCIWACIGLMAATKRDHNNYSKNPMRIVCYVGTWSVYHKVDPYTTIEDIDPFK 60
QY 61 CTHLMYGFAKIDYKYTIQVDFPYQDDNHNNSWEKRGYERFNNLRKKNPELTMTISLGWY 120
DB 61 CTHLMYGFAKIDYKYTIQVDFPYQDDNHNNSWEKRGYERFNNLRKKNPELTMTISLGWY 120
QY 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALV 180
DB 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALV 180
QY 181 RELKDAFEPHGYLLTAASVPGKIDRAYDIKELNKLFDWMVMTYDYHGGWENFYGHNA 240
DB 181 RELKDAFEPHGYLLTAASVPGKIDRAYDIKELNKLFDWMVMTYDYHGGWENFYGHNA 240
QY 241 PLYKRPDETDELHTYFNVNVTMYVYLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPK 300

DB 241 PLYKRPDETDELHTYFNVNVTMYVYLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPK 300
QY 301 KMSPPGISGEGVLSYIELCOLFQKEBWHIOYDEYVYNAFYNDKIWVGDDLASISCK 360
DB 301 KMSPPGISGEGVLSYIELCOLFQKEBWHIOYDEYVYNAFYNDKIWVGDDLASISCK 360
QY 361 KLAFLKELGVSGVWWSLENDDFKHCGRPNLLKVNHWINGDEKNSPECILGSPSTTPT 420
DB 361 KLAFLKELGVSGVWWSLENDDFKHCGRPNLLKVNHWINGDEKNSPECILGSPSTTPT 420
QY 421 TPTTPT 480
DB 421 TPTTPT 434
QY 481 TPTTPT 537
DB 435 ---PTTSTSETPKYTIIDGHLIKCYKQGYLPHPTDVHKYLVEYIATPNGGWWVHIM 491
QY 538 PCPCGTIWCQEKLTICGE 555
DB 492 DCPKGRWHTLKNKCIQE 509

RESULT 5

US-09-292-225-38
; Sequence 38, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-38

Query Match 81.8%; Score 2542; DB 4; Length 509;
Best Local Similarity 81.7%; Pred. No. 1.1e-195;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

QY 1 MKTIYALISIMACIGLMAASIKRDHNDYSKNPMRIVCYVGTWSVYHKVDPYTTIEDIDPFK 60
DB 1 MKTTFALFCIWACIGLMAATKRDHNNYSKNPMRIVCYVGTWSVYHKVDPYTTIEDIDPFK 60
QY 61 CTHLMYGFAKIDYKYTIQVDFPYQDDNHNNSWEKRGYERFNNLRKKNPELTMTISLGWY 120
DB 61 CTHLMYGFAKIDYKYTIQVDFPYQDDNHNNSWEKRGYERFNNLRKKNPELTMTISLGWY 120
QY 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALV 180
DB 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALV 180
QY 181 RELKDAFEPHGYLLTAASVPGKIDRAYDIKELNKLFDWMVMTYDYHGGWENFYGHNA 240
DB 181 RELKDAFEPHGYLLTAASVPGKIDRAYDIKELNKLFDWMVMTYDYHGGWENFYGHNA 240
QY 241 PLYKRPDETDELHTYFNVNVTMYVYLNNGATDKLVGMVFPFYGRAWSIEDRSKLGDPK 300

Db 241 PLYKPPDEDELTHTFNNVYTHYLLNNGATRDKLWGVFFYGRWSIEDRSKVLGDP 300
QY 301 KQSPPGFISGEGVLSYIELCOLFQKEWHIOYDEYNAPYNDKIWGYDDLASISC 360
Db 301 KQSPPGFISGEGVLSYIELCOLFQKEWHIOYDEYNAPYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSGVWVMSLENDDFKGCQKPNPLNKNVNMINGDEKNSFEICILQPSSTTP 420
Db 361 KLAFLKELGVSGVWVMSLENDDFKGCQKPNPLNKNVNMINGDEKNSFEICILQPSSTTP 420
QY 421 TPTTPT 480
Db 421 TPTTPT 434
QY 481 TSPPTTHTSTPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVEFV---NGGWWHIM 537
Db 435 ---PTTDDSTSTPKYTYVDGHLIKCYKQGYLPHTDVHKYLVECIATPNGGWWHIM 491
QY 538 PCPPGTWCQKLTICGE 555
Db 492 DCPKGRWHATLKNKICQE 509

RESULT 6
US-09-292-225-41
; Sequence 41, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-41

Query Match 79.7%; Score 2475; DB 4; Length 490;
Best Local Similarity 82.5%; Pred. No. 2.5e-190;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
QY 22 KRHDNDYKSNPRIVCYVGTWSVYHKVDPTYIEDIDPKCTHLMYGFADIDYKTYIQVF 81
Db 3 KRHDNDYKSNPRIVCYVGTWSVYHKVDPTYIEDIDPKCTHLMYGFADIDYKTYIQVF 62
QY 82 DPYQDDNHNSEKRYERFNNLRKNPBLTMSISLGWYEGSEKYSMDAANPTVYRQPFQ 141
Db 63 DPYQDDNHNSEKRYERFNNLRKNPBLTMSISLGWYEGSEKYSMDAANPTVYRQPFQ 122
QY 142 SVLDFLOEYKFDGLDWEYPCSRIGNPKIDKQNYLALVRELKDAFEHGYLLTAASPG 201
Db 123 SVLDFLOEYKFDGLDWEYPCSRIGNPKIDKQNYLALVRELKDAFEHGYLLTAASPG 182
QY 202 KDKIDRAYDIEKLNKLFDMWNVMTYDYGWENFYGHNAPLYKRPDETDLHTTFNNVYT 261
Db 183 KDKIDRAYDIEKLNKLFDMWNVMTYDYGWENFYGHNAPLYKRPDETDLHTTFNNVYT 242

QY 262 MHYLLNNGATRDKLWGVFFYGRWSIEDRSKVLGDPKAGMSPPGFISGEGVLSYIEL 321
Db 243 MHYLLNNGATRDKLWGVFFYGRWSIEDRSKVLGDPKAGMSPPGFISGEGVLSYIEL 302
QY 322 COLFQKEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSGVWVMSLEND 381
Db 303 COLFQKEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSGVWVMSLEND 362
QY 382 DFKGHCQKPNPLNKNVNMINGDEKNSFEICILQPSSTTPPTPTPTPTPTPTPTPTPTPT 441
Db 363 DFKGHCQKPNPLNKNVNMINGDEKNSFEICILQPSSTTPPTPTPTPTPTPTPTPTPTPT 415
QY 442 TPTTPT 501
Db 416 ---PTTDDSTSTPKYTYVD 433
QY 502 GHLIKCYKEGDIHPHTNIHKYLVEFV---NGGWWHIMPCPPGTWCQKLTICGE 555
Db 434 GHLIKCYKQGYLPHTDVHKYLVECIATPNGGWWHIMDCPKGRWHATLKNKICQE 490

RESULT 7

US-08-524-051-2
; Sequence 2, Application US/08524051
; Patent No. 5866788
; GENERAL INFORMATION:
; APPLICANT: Kramer, Karl J.
; APPLICANT: Muthukrishnan, Subbaratnam
; APPLICANT: Choi, Hee Kyung
; APPLICANT: Corpuz, Lolita
; APPLICANT: Gopalakrishnan, Bhuvana
; TITLE OF INVENTION: RECOMBINANT CHITINASE AND USE THEREOF AS
; TITLE OF INVENTION: A BIOCIDES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Blvd., Suite 400
; CITY: Kansas City
; STATE: MO
; COUNTRY: USA
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/524,051
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 22875-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (816)474-9050
; TELEFAX: (816)474-9057
; TELEX: 434-363
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-524-051-2

Query Match 27.0%; Score 838.5; DB 2; Length 554;
Best Local Similarity 35.4%; Pred. No. 4.8e-59;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;
QY 6 AILSIVACIGLWNASIKRDNDYKSNPRIVCYVGTWSVYHK-VDPYTIEDIDPKCTH 64
Db 3 ATLATLAVLATATV-----QSDRRARIVCYFNNWVYRFGVRYGIEDIPVEKCTHI 55

QY 65 MYGFAKIDYKYTIQVDFPDYQDDNHNHNSWEKRGYERFNNLRKLNPELTJTWISLGGWYEGSE 124
DB 56 IYSGFVTEGNSVLIIDPELD-----VDKNGFRNTSLRSHPSVKFWVAVGGWAGBS 110
QY 125 KYSDMAANPTYROOFIQSVLDFLQYKFDGLDLDWEYPGSR-LGNPKIDKONYLALVREL 183
DB 111 KYSHWAKSTRMSFIRSVSLKLYKDYDFDGLDLDWEYPGAADRGGSFSDKDFLYLVQEL 170
QY 184 KDAP--EPHGYSLLTAAPVSPGKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNA 241
DB 171 RRAFIRVGKWEWLTAAVPLANFRMLWEGHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDETDELHTYFNNVYTHYYLNGATRDKLVMGVPPFYGRAWSIED----- 290
DB 231 LYKRPD--QWYEXKLVNDGLHWEKGCPSNKLIVGIPFYGRSFTLSAGNNYGLGTPI 289
QY 291 RSKLGLGDPKAGMSPGPFISGREGVLSYELCOLFOKEE--WHIOYDEYNNAPYGYNDKI 348
DB 290 NKEAGGDPAPYTNATG-----WAYYEICTEVDKDDSGTWKWDGCKPCPYAYKGTQ 342
QY 349 WGYDDLASISCKLAFELKELGVSMVWSLENDDFKGCPCNPLLNKVNMMINGDEKNS 408
DB 343 WGYEDPRSVEIKWNWIKQKGYLGAMTWAIDMDDFQGLCGEKNPLIKILHKMS----- 396
QY 409 FECILGPSTTTPTTPTTPT-----TPTTPPS-----PTTPTTSPSTPTTTPSP 456
DB 397 -----SYTVPPPHENTTPTPEWARPPSPSDPSGDPPTTTTAKPASTTKTKTKT 449
QY 457 TPTTPTTSPIT-----PTP-TTPTAPTSTPSPTTHTSETPKYTYVVDGHLI 505
DB 450 TTTAKPQSVIDEENDINVRPEKPEPEVEVP-PTENE-----VDGSEI 497
QY 506 KYCEKGD-IPHTNIHKYLVCEVNGVGMVWHIMPCPPGIWCOEKLTC 552
DB 498 -CNSDQDYIPDKKXCDKYWC--VNGE--AMQFSCQHGTVFVVELNVC 540

RESULT 8
US-09-552-778-16
; Sequence 16, Application US/09052778A
; Patent No. 6060590
; GENERAL INFORMATION:
; APPLICANT: Bryant, Peter J.
; APPLICANT: Kawamura, Kazuo
; TITLE OF INVENTION: CHITINASE RELATED PROTEINS AND METHODS
; FILE REFERENCE: OF USE
; CURRENT APPLICATION NUMBER: 07306/015001
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Manduca sexta
US-09-552-778-16

Query Match 27.0%; Score 838.5; DB 3; Length 554;
Best Local Similarity 35.4%; Pred. No. 4.8e-59;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;
QY 6 AILSIWACIGLMAVSIKRDHNDYSKNPMRIVCYVGTWSVYHK-VDPYTIETDIPPKCTHL 64
DB 3 ATLATLAVLALATAV-----QSDSRARIVCFVSNWVYRPGVGRYGIETDIPVEKCTHI 55
QY 65 MYGFAKIDYKYTIQVDFPDYQDDNHNHNSWEKRGYERFNNLRKLNPELTJTWISLGGWYEGSE 124
DB 56 IYSGFVTEGNSVLIIDPELD-----VDKNGFRNTSLRSHPSVKFWVAVGGWAGBS 110
QY 125 KYSDMAANPTYROOFIQSVLDFLQYKFDGLDLDWEYPGSR-LGNPKIDKONYLALVREL 183
DB 111 KYSHWAKSTRMSFIRSVSLKLYKDYDFDGLDLDWEYPGAADRGGSFSDKDFLYLVQEL 170

QY 184 KDAP--EPHGYSLLTAAPVSPGKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNA 241
DB 171 RRAFIRVGKWEWLTAAVPLANFRMLWEGHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDETDELHTYFNNVYTHYYLNGATRDKLVMGVPPFYGRAWSIED----- 290
DB 231 LYKRPD--QWYEXKLVNDGLHWEKGCPSNKLIVGIPFYGRSFTLSAGNNYGLGTPI 289
QY 291 RSKLGLGDPKAGMSPGPFISGREGVLSYELCOLFOKEE--WHIOYDEYNNAPYGYNDKI 348
DB 290 NKEAGGDPAPYTNATG-----WAYYEICTEVDKDDSGTWKWDGCKPCPYAYKGTQ 342
QY 349 WGYDDLASISCKLAFELKELGVSMVWSLENDDFKGCPCNPLLNKVNMMINGDEKNS 408
DB 343 WGYEDPRSVEIKWNWIKQKGYLGAMTWAIDMDDFQGLCGEKNPLIKILHKMS----- 396
QY 409 FECILGPSTTTPTTPTTPT-----TPTTPPS-----PTTPTTSPSTPTTTPSP 456
DB 397 -----SYTVPPPHENTTPTPEWARPPSPSDPSGDPPTTTTAKPASTTKTKTKT 449
QY 457 TPTTPTTSPIT-----PTP-TTPTAPTSTPSPTTHTSETPKYTYVVDGHLI 505
DB 450 TTTAKPQSVIDEENDINVRPEKPEPEVEVP-PTENE-----VDGSEI 497
QY 506 KYCEKGD-IPHTNIHKYLVCEVNGVGMVWHIMPCPPGIWCOEKLTC 552
DB 498 -CNSDQDYIPDKKXCDKYWC--VNGE--AMQFSCQHGTVFVVELNVC 540

RESULT 9
US-09-545-814-2
; Sequence 2, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-2

Query Match 26.2%; Score 815.5; DB 4; Length 583;
Best Local Similarity 34.6%; Pred. No. 3.6e-57;
Matches 198; Conservative 94; Mismatches 205; Indels 75; Gaps 23;
QY 7 ILSIWACIGLMAVSIKRDHNDYSKNPMRIVCYVGTWSVYHK-VDPYTIETDIPPKCTHL 65
DB 6 LLAVLCALAISSINTV-----EASDQKARIVCFVSNWVYRPGVGRYGIETDIPVLCITIV 61
QY 66 YGFAKIDYKYTIQVDFPDYQDDNHNHNSWEKRGYERFNNLRKLNPELTJTWISLGGWYEGSE 125
DB 62 YSFIGVDDKQSVLVIDPELDIDN-----GPKNFTNLRKTHPNVQLQIAGVGGWAGGKK 116
QY 126 YSDMAANPTYROOFIQSVLDFLQYKFDGLDLDWEYPGSR-LGNPKIDKONYLALVREL 184
DB 117 YSTWVAERKRSAFIRSVDFMVEYKFDGDFLDWEYFGAADRGGSFSDKDFLYLVQEL 176
QY 185 DAFEPHG--YLLTAAPVSPGKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNA 242
DB 177 RAFNKGQKWEITWAVPIAKFRLQEGVHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 236
QY 243 YKRPDETDELHTYFNNVYTHYYLNGATRDKLVMGVPPFYGRAWSIEDRSK-LKLG--D 298

Db 237 YRRPHD-QYAYEKLNVNDGLQLWDMGCPANKLVGVFPYGRSFTLSNKNKDYRLGTIN 295
QY 299 PAKGMSPPGFTSGEGVLSTYELCOLFO--KEEWHIOYDEYNAPYGYNDKIWGVYDCLA 356
Db 296 KEAGGEPGPTNATGFTSYEIECLEVDPSKGTWKWDEHGKVPYAYKQNWGVYEDPK 355
QY 357 SISCKLAPLKLGVSGVMVNSLENDPFKGHC--PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 356 SVALKMEFIKSGYGGAMTWMDDDFQVCSDDKHTLAVIMHDY---KNYIVPEFDS 411
QY 416 STTTPTP--TTTPTPTPTPTPTPTPT--TPSPT--TPPTPTPTPTPTPTPTPTPTPTPT 471
Db 412 SRTTPREWAKPPTPSQEP--DDTPYIPTTHAPKSRKPKPKPT--TTTAAATTPVAT 468
QY 472 TPTAPTSTPSPTTE-----HTSETP-----KYTTYVDGHLKCYKEGDIPHT 517
Db 469 -----TTTEHHHHHEEKSEBQDNQVGSQDTATD---VDCSQEDYLPH-E 511
QY 518 NIHKYLVCFVNGGWWH-----IMPCPPGTIW 545
Db 512 DCKNYRC-----VHGEAVLFTCREGTVY 535
RESULT 10
US-09-545-814-5
; Sequence 5, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-5
Query Match 26.2%; Score 815.5; DB 4; Length 583;
Best Local Similarity 34.6%; Pred. No. 3.6e-57;
Matches 198; Conservative 94; Mismatches 205; Indels 75; Gaps 23;
QY 7 ILSIWACIGLWNASIKRDNDYKKNPMRIVCYVGTWVSYHK-VDPYTIEDIDPFKCTHLM 65
Db 6 LLAVLCAIAISSINTV----EASDQKARIVCYFSNVAVTRPGIGRYGIEDIPVDLCTHIV 61
QY 66 YGPAKIDYKYTIQVDFPYQDDNHSWEKRGYERPNLRLKNPELTMTWISLGGWYEGSEK 125
Db 62 YSFIGVDDKWSVLVIDPELDIDN----GFKNFTNLRKHHPNVKLQIAGVGAEGGKK 116
QY 126 YSDMAANPTYRQOFTQSVLDFLOEYKFDGLDWEYPGSR--LGNPKIDKQNYLALVRELK 184
Db 117 YSTWVAEKRKSAFIRSVVDFMNEYKFDGLDWEYPGAADRGSGFSKDKFLYFVQELR 176
QY 185 DAFEPHG--YLLTAAVSPGKKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAFL 242
Db 177 RAFNKGQKNWEITWAVPTAKFRLQEGYHVPELCELLDAIHVMSYDLRGNWAGFADTHSPL 236
QY 243 YKRDETDDELTYFNWNTMYHLYNNGATRKLVNGVDFPYGRAWSIEDRSK--LKLG---D 298
Db 237 YRRPHD-QYAYEKLNVNDGLQLWDMGCPANKLVGVFPYGRSFTLSNKNKDYRLGTIN 295
QY 299 PAKGMSPPGFTSGEGVLSTYELCOLFO--KEEWHIOYDEYNAPYGYNDKIWGVYDCLA 356
Db 296 KEAGGEPGPTNATGFTSYEIECLEVDPSKGTWKWDEHGKVPYAYKQNWGVYEDPK 355
QY 357 SISCKLAPLKLGVSGVMVNSLENDPFKGHC--PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 416 STTTPTP--TTTPTPTPTPTPTPTPT--TPSPT--TPPTPTPTPTPTPTPTPTPTPTPT 471
QY 472 TPTAPTSTPSPTTE-----HTSETP-----KYTTYVDGHLKCYKEGDIPHT 517
Db 469 -----TTTEHHHHHEEKSEBQDNQVGSQDTATD---VDCSQEDYLPH-E 511
QY 518 NIHKYLVCFVNGGWWH-----IMPCPPGTIW 545
Db 512 DCKNYRC-----VHGEAVLFTCREGTVY 535
RESULT 11
US-09-545-814-32
; Sequence 32, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 32
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Tagged Ctenocephalides felis
US-09-545-814-32
Query Match 26.2%; Score 812.5; DB 4; Length 635;
Best Local Similarity 34.6%; Pred. No. 7e-57;
Matches 198; Conservative 93; Mismatches 206; Indels 75; Gaps 23;
QY 7 ILSIWACIGLWNASIKRDNDYKKNPMRIVCYVGTWVSYHK-VDPYTIEDIDPFKCTHLM 65
Db 7 LLAVLCAIAISSINTV----EASDQKARIVCYFSNVAVTRPGIGRYGIEDIPVDLCTHIV 62
QY 66 YGPAKIDYKYTIQVDFPYQDDNHSWEKRGYERPNLRLKNPELTMTWISLGGWYEGSEK 125
Db 63 YSFIGVDDKWSVLVIDPELDIDN----GFKNFTNLRKHHPNVKLQIAGVGAEGGKK 117
QY 126 YSDMAANPTYRQOFTQSVLDFLOEYKFDGLDWEYPGSR--LGNPKIDKQNYLALVRELK 184
Db 118 YSTWVAEKRKSAFIRSVVDFMNEYKFDGLDWEYPGAADRGSGFSKDKFLYFVQELR 177
QY 185 DAFEPHG--YLLTAAVSPGKKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAFL 242
Db 178 RAFNKGQKNWEITWAVPTAKFRLQEGYHVPELCELLDAIHVMSYDLRGNWAGFADTHSPL 237
QY 243 YKRDETDDELTYFNWNTMYHLYNNGATRKLVNGVDFPYGRAWSIEDRSK--LKLG---D 298
Db 238 YRRPHD-QYAYEKLNVNDGLQLWDMGCPANKLVGVFPYGRFFTLSNKNKDYRLGTIN 296
QY 299 PAKGMSPPGFTSGEGVLSTYELCOLFO--KEEWHIOYDEYNAPYGYNDKIWGVYDCLA 356
Db 297 KEAGGEPGPTNATGFTSYEIECLEVDPSKGTWKWDEHGKVPYAYKQNWGVYEDPK 356
QY 357 SISCKLAPLKLGVSGVMVNSLENDPFKGHC--PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 357 SVALKMEFIKSGYGGAMTWMDDDFQVCSDDKHTLAVIMHDY---KNYIVPEFDS 412
QY 416 STTTPTP--TTTPTPTPTPTPTPTPT--TPSPT--TPPTPTPTPTPTPTPTPTPTPTPT 471
Db 413 SRTTPREWAKPPTPSQEP--DDTPYIPTTHAPKSRKPKPKPT--TTTAAATTPVAT 469

Db 237 YRRPHD-QYAYEKLNVNDGLQLWDMGCPANKLVGVFPYGRSFTLSNKNKDYRLGTIN 295
QY 299 PAKGMSPPGFTSGEGVLSTYELCOLFO--KEEWHIOYDEYNAPYGYNDKIWGVYDCLA 356
Db 296 KEAGGEPGPTNATGFTSYEIECLEVDPSKGTWKWDEHGKVPYAYKQNWGVYEDPK 355
QY 357 SISCKLAPLKLGVSGVMVNSLENDPFKGHC--PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 356 SVALKMEFIKSGYGGAMTWMDDDFQVCSDDKHTLAVIMHDY---KNYIVPEFDS 411
QY 416 STTTPTP--TTTPTPTPTPTPTPTPT--TPSPT--TPPTPTPTPTPTPTPTPTPTPTPT 471
Db 412 SRTTPREWAKPPTPSQEP--DDTPYIPTTHAPKSRKPKPKPT--TTTAAATTPVAT 468
QY 472 TPTAPTSTPSPTTE-----HTSETP-----KYTTYVDGHLKCYKEGDIPHT 517
Db 469 -----TTTEHHHHHEEKSEBQDNQVGSQDTATD---VDCSQEDYLPH-E 511
QY 518 NIHKYLVCFVNGGWWH-----IMPCPPGTIW 545
Db 512 DCKNYRC-----VHGEAVLFTCREGTVY 535
RESULT 10
US-09-545-814-5
; Sequence 5, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-5
Query Match 26.2%; Score 815.5; DB 4; Length 583;
Best Local Similarity 34.6%; Pred. No. 3.6e-57;
Matches 198; Conservative 94; Mismatches 205; Indels 75; Gaps 23;
QY 7 ILSIWACIGLWNASIKRDNDYKKNPMRIVCYVGTWVSYHK-VDPYTIEDIDPFKCTHLM 65
Db 6 LLAVLCAIAISSINTV----EASDQKARIVCYFSNVAVTRPGIGRYGIEDIPVDLCTHIV 61
QY 66 YGPAKIDYKYTIQVDFPYQDDNHSWEKRGYERPNLRLKNPELTMTWISLGGWYEGSEK 125
Db 62 YSFIGVDDKWSVLVIDPELDIDN----GFKNFTNLRKHHPNVKLQIAGVGAEGGKK 116
QY 126 YSDMAANPTYRQOFTQSVLDFLOEYKFDGLDWEYPGSR--LGNPKIDKQNYLALVRELK 184
Db 117 YSTWVAEKRKSAFIRSVVDFMNEYKFDGLDWEYPGAADRGSGFSKDKFLYFVQELR 176
QY 185 DAFEPHG--YLLTAAVSPGKKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAFL 242
Db 177 RAFNKGQKNWEITWAVPTAKFRLQEGYHVPELCELLDAIHVMSYDLRGNWAGFADTHSPL 236
QY 243 YKRDETDDELTYFNWNTMYHLYNNGATRKLVNGVDFPYGRAWSIEDRSK--LKLG---D 298
Db 237 YRRPHD-QYAYEKLNVNDGLQLWDMGCPANKLVGVFPYGRSFTLSNKNKDYRLGTIN 295
QY 299 PAKGMSPPGFTSGEGVLSTYELCOLFO--KEEWHIOYDEYNAPYGYNDKIWGVYDCLA 356
Db 296 KEAGGEPGPTNATGFTSYEIECLEVDPSKGTWKWDEHGKVPYAYKQNWGVYEDPK 355
QY 357 SISCKLAPLKLGVSGVMVNSLENDPFKGHC--PKNPLLNKVNMMINGDEKNSFECILGP 415

QY 472 TPTPATTSTRPTTTE-----HTSETP-----KYTTVDGHLIKYKGGDIHPT 517
Db 470 -----TTTEHHHHHEEKSEQDNQVGSQDTATD---VDCSEDYLPB-E 512
QY 518 NIHKLVCFFVNGGWWVH-----IMPCPPGTIW 545
Db 513 DCKKYRC-----VHGEAVLFTCRBGTVY 536

RESULT 12
US-09-545-814-14
; Sequence 14, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE REFERENCE: PC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-14

Query Match 26.1%; Score 811.5; DB 4; Length 559;
Best Local Similarity 35.8%; Pred. No. 7.1e-57;
Matches 195; Conservative 86; Mismatches 193; Indels 71; Gaps 22;

QY 34 RIVCYGVTSVYHK-VDPYTTIEDIDPFKCTHLMYGAFAKIDYKTYIQVDPYQDDNHSW 92
Db 5 RIVCYFSNWAYRPGIRGYGIEDIPVLCITHIVSFIGVDDKMSVLVIDPDLIDDN-- 62
QY 93 EKRGVYERFNNLRKMPETLMTSLGWEYEGSEKYSMDAANTPYRQOFTQSVLDLFLQEVKF 152
Db 63 ---GFKFNLRKIHFNKVLQAVGNAEGGKYSTVAEKKRSASFIRSVVDFMNEYKF 119
QY 153 DGLDLWDWEYVPSGR-LGNPKIDKQNTLALVRELKDAFEPHG--YLLTAAVSPGKKIDRAY 209
Db 120 DGFDLWDWEYVPGADRGGSEFSDKXFLYFQELRRAFNKQGNWELTMAVPIAKFLQEGY 179
QY 210 DTKELNKLFDWMNVMTYDYGWENFYGHNAFLYKRPDETDELHTYFNVTMHTYLNNG 269
Db 180 HVPCLCELLDAIHVMSYDLRGNWAGFADTHSPLYRRPHD-QVAYEKLNVNDGLQLWVDMG 238
QY 270 ATRDKLVGVPPYGRAWSIEDRSK-LKLG---DPAKGMSPGPGFISGEGVLSYIELCOLF 325
Db 239 CPANKLVGVPPYGRSFTLSNSNDYRLGTINKEAGGEGPGPYTNATGTSYVEICLEV 298
QY 326 Q--KEEMHIQVDEYNAPYGVNDKLVGVYDDLASISCKLAFKLKELGVGMVWVSLNDDF 383
Db 299 DDPKSGWTKKDEHOKVPYAKGNQVGYEDPKSVALKQEFIKSGYGMATWAIMDDF 358
QY 384 KGHCG-PKNPILNKVHMMINGDEKNSPECILGPSFTTPTP--TTTPTPTPTTPTTPTT 440
Db 359 QGVCSDDKXHTLAVIMHDYK---KNYIVPEPDSRITRPERWAKPPTSPQEP-DDTPIYI 413
QY 441 PTT-TPSPPT-PTTTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 493
Db 414 PTHAPKPSRKPTRKPKPT--TTTVAATTVPAT-----TTTEHHHHHEEEK 458
QY 494 P-----KYTTVDGHLIKYKGGDIHPTNIHKLVCFFVNGGWWVH-----IMPCP 540
Db 459 PSEQDNQVGSQDTATD---VDCSEDYLPB-EDCKNYRC-----VHGEAVLFTCR 506
QY 541 PGTIW 545
Db 507 EGTIV 511

RESULT 13
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-486-839-4

Query Match 23.8%; Score 738; DB 2; Length 466;
Best Local Similarity 34.2%; Pred. No. 4.3e-51;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;

QY 34 RIVCYGVTSVYHKVDP-YTIEDIDPFKCTHLMYGAFAKIDYKTYIQVDPYQDDNHSW 92
Db 23 KLVCFYFTNMAQYRQGEARFLPKDLPDLSCLTHTLYAFAGMTNHLST-----TEW 71
QY 93 -EKRGVYERFNNLRKMPETLMTSLGWEYEGSEKYSMDAANTPYRQOFTQSVLDLFLQEVK 151
Db 72 NDTLLYQEFNGLKQKMPKLTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYS 131
QY 152 FDGLDLWDWEYVPSGR-LGNPKIDKQNTLALVRELKDAFEPHG-----YLLTAAVSPGKK 204
Db 132 FDGLDLWDWEYVPSGQ-OSPVDKXERFTLLYQDLANAFQQAQTSKRELJLSAARVAGQY 190
QY 205 IDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAFLYKRPDETDELHTYFNVTMHTY 264
Db 191 VDAGYEVDKIAQNLDVNLMAVDFHGSWEKVTGHNSPLYKQESGAAS-LNVDAVQ 249
QY 265 YLANGATRDKLWGVFPYGRAWSIEDRSKLGDPKAGMSPGPGFISGEGVLSYIELCOL 324
Db 250 WLQKGTPTASKLILGMPTCYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYVEVCS 309
QY 325 FQKEWHIOVDEYNAPYGVNDKLVGVYDDLASISCKLAFKLKELGVGMVWVSLNDDFK 384
Db 310 KGATKORIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSYLKQKGLGGVWVWALDDFA 366

Search completed: March 22, 2004, 07:03:58
Job time : 45.7332 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:40:28 ; Search time 36.0951 Seconds
(without alignments)
1479.047 Million cell updates/sec

```

Title: US-09-662-293-18
Perfect score: 3107
Sequence: 1 MKTIYAILSLINACIGLMNAS.....IMCPPGTIWCQEKLTGICE 555

```

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2833366

Minimum DB seq length: 0

Maximum DB per length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:★

```
1: girl:*
```

```
2: pir2:*
```

3: pir3:*

4: pi4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|-----------------------|
| 1 | 1008.5 | 32.5 | 525 | 2 | T44445 | chitinase (EC 3.2.2.) |
| 2 | 869 | 28.0 | 1635 | 2 | T14075 | chitinase (EC 3.2.2.) |
| 3 | 838.5 | 27.0 | 554 | 2 | A55596 | chitinase (EC 3.2.2.) |
| 4 | 802 | 25.8 | 483 | 2 | A53918 | chitinase (EC 3.2.2.) |
| 5 | 705.5 | 22.7 | 617 | 2 | T13408 | hypothetical prote |
| 6 | 685.5 | 22.1 | 537 | 2 | S51797 | oviduct-specific g |
| 7 | 685 | 22.0 | 539 | 2 | I46470 | estrogen dependent |
| 8 | 668 | 21.5 | 654 | 2 | I38605 | oviductal glycopro |
| 9 | 662 | 21.3 | 383 | 2 | A49562 | cartilage glycopro |
| 10 | 650 | 20.9 | 383 | 2 | S51327 | heparin-binding gl |
| 11 | 646.5 | 20.8 | 405 | 2 | S61551 | breast-regressing |
| 12 | 608 | 19.6 | 399 | 2 | S27879 | secretory protease |
| 13 | 604.5 | 19.5 | 504 | 2 | A38221 | chitinase (EC 3.2.2.) |
| 14 | 592 | 19.1 | 599 | 2 | D87364 | chitinase BHO916 [|
| 15 | 561 | 18.1 | 699 | 2 | A38368 | chitinase EC 3.2.2. |
| 16 | 560 | 18.0 | 1315 | 2 | T43916 | chitinase A (impor |
| 17 | 463.5 | 14.9 | 831 | 2 | T00323 | chitinase (EC 3.2.2.) |
| 18 | 432.5 | 13.9 | 1484 | 2 | T23275 | hypothetical prote |
| 19 | 428 | 13.8 | 869 | 2 | T44440 | chitinase EC 3.2.2. |
| 20 | 427.5 | 13.8 | 849 | 2 | D82510 | chitinase VCA0027 |
| 21 | 424.5 | 13.7 | 2025 | 2 | T03884 | hypothetical prote |
| 22 | 414.5 | 13.3 | 563 | 2 | S60651 | chitinase precursor |
| 23 | 413 | 13.3 | 423 | 2 | A14552 | chitinase B homolo |
| 24 | 412 | 13.3 | 756 | 2 | QJ1975 | chitinase (EC 3.2.2.) |
| 25 | 411.5 | 13.2 | 424 | 2 | S68121 | chitinase I precur |
| 26 | 411 | 13.2 | 423 | 2 | S53129 | chitinase - fungu |
| 27 | 409.5 | 13.2 | 756 | 2 | AA1088 | chitinase B homolo |
| 28 | 408.5 | 13.1 | 561 | 2 | T25090 | chitinase (EC 3.2.2.) |
| 29 | 405 | 13.0 | 546 | 2 | F84328 | chitinase [impor |
| 30 | 402 | 13.0 | 546 | 2 | A24238 | chitinase [impor |

| | | | | | |
|----|-------|------|------|---|--------|
| 30 | 394 | 12.7 | 452 | 2 | JC4038 |
| 31 | 394 | 12.7 | 633 | 2 | T24898 |
| 32 | 391.5 | 12.6 | 558 | 2 | T30418 |
| 33 | 389 | 12.5 | 424 | 2 | S47133 |
| 34 | 386 | 12.5 | 820 | 2 | A40633 |
| 35 | 385 | 12.4 | 799 | 2 | PC4106 |
| 36 | 382 | 12.3 | 755 | 2 | T20950 |
| 37 | 379.5 | 12.2 | 765 | 2 | T35719 |
| 38 | 377.5 | 12.1 | 550 | 2 | T10393 |
| 39 | 371.5 | 12.0 | 398 | 2 | T04761 |
| 40 | 369 | 11.9 | 1088 | 2 | D82246 |
| 41 | 367.5 | 11.8 | 511 | 2 | S61166 |
| 42 | 365 | 11.7 | 379 | 2 | T04762 |
| 43 | 364.5 | 11.7 | 552 | 2 | T41863 |
| 44 | 360.5 | 11.6 | 551 | 2 | G72865 |
| 45 | 359.5 | 11.6 | 427 | 2 | JC4565 |

ALIGNMENTS

RESULT 1

REF ID: A64445

chitinase (EC 3.2.1.14) [imported] - African malaria mosquito

C/Species: Anopheles gambiae (African malaria mosquito)

```
C;Date: 21-Jan-2000 #sequence revision 21-Jan-2000 #text change 21-Jan-2000
```

C;Accession: T44445

R;Shen, Z.; Jacobs-Lorena, M.

submitted to the EMBL Data Library, June 1997

A;Reference number:

A;Accession: T44445

A;Status: preliminary;

A;Molecule type: mRNA

A;Residues: 1-525 <SHE>

A;Cross-references: EMBL:AF008575

A; Experimental source: adult; gut

C;Genetics:

A;Gene: chi-1

C;Keywords: glycosidase; hydrolase

L
C
L

J
J
I
C
T

C
D
P

L
C
C
C
T

C
J
C
U
C

J
d
L
C
C

T
t
M
S
C
C
C

Query Match 32.5%; Score 1008.5; DB

Best Local Similarity 38.4%; Pred. No. 2.2e-55;
Matches 211; Concentrations 23; Wavelengths 207

[illegible]

Db 414 APTT-----STVAPGTTTTTGANPGTTPPT--SDAPNHTTTTTEGNGPGRTPPSG 466
Qy 485 TTEHTSETPKYTVVDGHLKCY--KEGDIPHPNIHKYLCV-----EFVNGGWWVH 535
Db 467 -----DG---PCAGGRIGFVHPINCARYYICLTADTYIEFT----- 500
Qy 536 IMPCPPGRIW 545
Db 501 ---CPPGTLF 507
RESULT 2
T14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14075
R:de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A:Reference number: Z17872
A:Accession: T14075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1635
A:Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AAB81850.1
C:Genetics:
A:Gene: CHT2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 28.0%; Score 869; DB 2; Length 1635;
Best Local Similarity 34.3%; Pred. No. 4.8e-46;
Matches 196; Conservative 95; Mismatches 196; Indels 84; Gaps 18;
Qy 21 IKRDNDYSKNPM-----RIVCYVGTWSVHKVD-PYTIEDIDPFKCTHLMYFPAKIDE 73
Db 650 INKETSQDENAIESDVKVVCYTNWATYRQGNKYLPELIDADLCTHIVIGFAVLDR 709
Qy 74 YKVTIQVDFPYQDDNHNW---EKRGYERFNNLRNKPENLPTMTISLGGWYEGS-EKYSDM 129
Db 710 DRLVIK-----PHDSWADIDNFEYERVVEYKKGKVT--VAIGWNDSAGDKYSLR 759
Qy 130 ANPTYRQFIQSVLDFLOEYKFDGLDLDWEYVPSRLGNPK-----IDKNYLALVRELKD 185
Db 760 VRSAAARQKFIADVAFIEKYFGDGLDLDWEYVPCWQVDCCKGFSDEKEGFASLIVELSQ 819
Qy 186 AFEPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKR 245
Db 820 AFKPKGLLSAVSPSKVVDGYDVVTLSDYMDIAVNAVYDYGQWDKKTGHVAPMYEH 879
Qy 246 PDETDELHTYFNNVTMYLLANGATRDKLVMGVFPFYGRANSIEDRSKLGLDPAGKMS 305
Db 880 PDDFKT---FNAFTIHWIEKGDADPKLVGMGPMYQSGFSLANKHEGLNAKTYGGGE 936
Qy 306 PGFISGEGVLSYIELCOLFOKEWHIOYDEYN-APYGYNDKIWVGYDDLASISCKLAF 364
Db 937 AGESTRARGFLSYEICANIRKNKWTAVDRKRGVPYAKGDQVSPDDQVMIRHKEY 996
Qy 365 LKELGVSGVWMSLENDPFGHCG-PKNPLNKVHNMINGDEKNSFECIL--GPSTTTPT 421
Db 997 VKAMGLGGAMWALDLDLDFNLCDCEYPLLRINVLNYPGPRCVLEKEPQREBP 1056
Qy 422 PTTTPTTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 481
Db 1057 PTPRPTS-----TPPTTRPSTTISRRTTMTTIT--TTTRPTTTTTRRTSAR 1104
Qy 482 PSPTTTHTTSETPKY-----TYVDGHLIKCVKEGDIHPHTNIHKYLVCFEFGWG 532
Db 1105 PVTYIT---NVPSYQEIANEVDEFTCTDGRLP-----VPHPTDCKNYICQYKG--- 1150
Qy 533 WVHIMPCCPGTIW-----CQEKLT 551

Db 1151 -----LCPGLYWSVDHCDWPOSTNCRNKQT 1176
RESULT 3
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C:Accession: A56596
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitin
A:Reference number: A56596; MUID:93357793; PMID:8353525
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:U02270; GB:S64757; NID:g406048; PID:g406049
A:Experimental source: larvae
A:Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIP:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 27.0%; Score 838.5; DB 2; Length 554;
Best Local Similarity 35.4%; Pred. No. 9.2e-45;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;
Qy 6 AILSINACIGLMMASIKRDNDYSKNPMELVCIVGTWTSVYHK-VDPYTIEDIDPFKCTHL 64
Db 3 ATLATLAVLALATAV-----QSDSRARIVCFNSVAVYRPGVGRYGIEDIPVEXCTHI 55
Qy 65 MYGFAKIDYKTIQVDFPYQDDNHNWSEKRGYERFNNLRNKPENLPTMTISLGGWYEGS 124
Db 56 IYFIFGVTGENSEVLIDPELD-----VDKNGFRNTSLRSSHPSVKFVWVAVGWAEGSS 110
Qy 125 KYSDMAANPTYRQFIQSVLDFLOEYKFDGLDLDWEYVPSGR-LGNPKIDKQNYLALVREL 183
Db 111 KYSHWVAQKSTRMSFIRSVVSLKKYDFDGLDLDWEYVPGADRGGSGFSKDKFLYLVLQEL 170
Qy 184 KDAF--EPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 241
Db 171 RRAFIRVKGWELTAAVPLANFRLMEGYHVPCLQELDAIHVMYSYDLRGNWAGFADVHSP 230
Qy 242 LYKRPDETDELHTYFNNVTMYLLANGATRDKLVMGVFPFYGEAWSIED----- 290
Db 231 LYKRPDH-QWAYBKLVNDGLHLWEEKGPCSNKLVVGIPIFYGRSFTLSAGNNNYGLGTFI 289
Qy 291 RSKLGLDPAGKMSPPGFISSGEGVLSYIELCOLFOKEE--WHIOYDEVYNAPYGYNDKI 348
Db 290 NKEAGGDPAPYTNATGP-----WAYVEICTEVKDDSGWTKKWDQCKCPYAKGTQ 342
Qy 349 WVGYYDDLASISCKLAFELKELGVSGVWMSLENDPFGHCGPKNPLNKVHNMINGDEKNS 408
Db 343 WVGVEDPSRVEIKWNWIKQKGLGAMTWMDDFQGLGCKGNPLIKILHKMS----- 396
Qy 409 FFCILGSPSTTTPTTPTTPTTPTT-----TTPTPS-----PTTPTTPTTPTTPTTPTT 456
Db 397 -----SYTVRPPHTENTTPEWARPPSTPDPSEGDPIPTTTTAKFASTTKTKVKT 449
Qy 457 TPTTPTSPPT-----PTP-TTPTAPTSTPTTPTTPTTPTTPTTPTTPTTPTTPTT 505
Db 450 TTTTAKPQSQVIDEENDINVRPEKPEPQPEPEVEVP-PTENE-----VDSGEI 497
Qy 506 KCYKEGG-IHPHTNIHKYLVCFEFGWGWVHIMPCCPGTIWCCQKLT 552
Db 498 -CNSDDQDIIPDKKHCKYKWC--VNGE--AMQFSCQHGIVFNVELNVC 540
RESULT 4
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C:Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000

C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and
A;Reference number: A49562; MUID:94084658; PMID:845017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:G348911; PIDN:AAA16074.1; PID:G348912
R;Nyirkos, P.; Golds, E.E.
Biochem. J. 269, 265-266, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary prote
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40, 'X', 42-45 <N2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 21.3%; Score 662; DB 2; Length 383;
Best Local Similarity 34.9%; Pred. No. 5.5e-34;
Matches 144; Conservative 81; Mismatches 136; Indels 52; Gaps 12;

QY 2 KTIYAILSIWACGLMNASIKRDNYSKNPMAIVCVGTWVYHKVDPTIED-IDPDK 60
DB 7 QTGFVLLVLLQCC-----SAYKLVCYTSWSQYREGDSCFPDLDREL 50
QY 61 CTHLMYGFADIKYKTIQVDPYQDDNHNHSEKRG--YERFNNRLKNPELTMTISLG 117
DB 51 CTHIYGFANI-----SNDHIDTWNDVTLGMNTLKNRNPNTLTLVSG 97
QY 118 GWTEGSESKDMAANPTYRQFIQSVDLDFQEKYFQGLDLDWEPYSGRLGNPKIDKQNYL 177
DB 98 GWNFGSQRFKSIASNTSRFTKSVPPFLRTGFGDLDAWLYPGER-----DKQHT 151
QY 178 ALVRELKDAF-----EP--HGVLTAAVSPGKIDRAYDIKELNKLFDMMNVNTYDHGG 231
DB 152 TLIKEMAEFIKEAQPGKQKLLSALSAGKVTIDSSYDIKTSQHLDFISIMYDFHGA 211
QY 232 WENFYGNAPLYKRPDTELDHLYFNVTYNNYLNAGTRDKLVGMVPPYGRAWSIEDR 291
DB 212 WRGTGTHSPFLF-RGQDASDPDRSNTDYAVGYMLRLGAPASKLVMGIPTFGRSFTLAS- 269
QY 292 SKLKGDPAGKMSPPGISBEGVLSYELCOLFQKEEMHIQYDEYNNAPYGYNDKIWVG 351
DB 270 SETGVGAPISGPGIPGRFTKEAGTLAYEICDFLRGATVHRTLGQ--QVPIATKGNQWVG 327
QY 352 YDDLASISCKLAFIKELGVSGVMWSLENDDFKG-HCGP--KNPLLKNKHMI 401
DB 328 YDDQESVKSQVYKDKRQLAGAMWALDLDLDFQSGFCQDLRPLTNIAKDAL 380

RESULT 10
S51327
heparin-binding glycoprotein 38K - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackleton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differ
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:Z47803; NID:9634097; PIDN:CAA87764.1; PID:9634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 20.9%; Score 650; DB 2; Length 383;

Best Local Similarity 36.6%; Pred. No. 3.1e-33;
Matches 139; Conservative 77; Mismatches 130; Indels 34; Gaps 11;

QY 34 RIVCVGTWVYHKVDPTIED-IDPDKCTHLMYGFADIKYKTIQVDPYQDDNHNH 91
DB 23 KLVCYTSWSQYREGDSCFPDAIDPFLCTHIIYSFANISNNEIDTLE----- 70
QY 92 W-EKRGYERFNNRLKNPELTMTISLGWYEGSKYSDMAANPTYRQFIQSVDLDFQ 150
DB 71 WNDVTLYTLNTLKNRNPNTLTLVSGWNFSGRPSKIANTQSRRTFKISVPFLTH 130
QY 151 KPDGLDLDWEPYSGRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTA 204
DB 131 GPDGLDAWISPGRR-----DKRHLLTVLKEMAEFVREALPGTERLLSGAVS 184
QY 205 IDRAYDIKELNKLFDMMNVNTYDHGGWENFYGNAPLYKRPDTELDHLYFN 264
DB 185 IDRGYDIAIQSHLDLFIQLTYDPHGAWRQTTHHSPLFRGQGDASS-DRFSN 243
QY 265 YLNNCATRDKLVGMVPPYGRAWSIEDRSLKLGDPAGKMSPPGISBEGVLSYEL 324
DB 244 VLRLGAPANKLVMGIPTFGRSFTLAS-SKTDVGAPASGPGIPGRFTKEGIL 302
QY 325 FQKEEWHIYQYDEYNNAPYGYNDKIWVGYYDDLASISCKLAFIKELGVSGVM 384
DB 303 LOGAT--VRPLGQGVATKGNQWVGYYDDQESVKNKAKYLKSRQLAGAMW 360
QY 385 GH-CGP--KNPLLKNKHMI 401
DB 361 GNFCQNLRFPLTSAIKDVL 380

RESULT 11
S61551
breast-regressing protein brp39 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S61551; S61550; I48271
R;Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A;Title: neu and ras initiate murine mammary tumors that share genetic markers general
A;Reference number: I48271; MUID:95060797; PMID:7970700
A;Accession: S61551
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-405 <MORI>
A;Cross-references: EMBL:X93035; NID:G1085065; PIDN:CAA63603.1; PID:G1085066
R;Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A;Reference number: S61550
A;Accession: S61550
A;Molecule type: mRNA
A;Residues: 1-245, 'I', 247-330, 'H', 332-350, 'MWALDLDLDFQGTCPQKPEFFPLTNIAKDALA', <MOR2>
A;Cross-references: EMBL:X93035; NID:G1085065; PIDN:CAA63603.1; PID:G1085066
A;Note: the differences at the carboxyl end are due to a frameshift error
C;Genetics:
A;Gene: brp39
C;Superfamily: Streptomyces chitinase chi40
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

Query Match 20.8%; Score 646.5; DB 2; Length 405;
Best Local Similarity 34.4%; Pred. No. 5.5e-33;
Matches 140; Conservative 73; Mismatches 129; Indels 65; Gaps 12;

QY 34 RIVCVGTWVYHK-VDPYTIEDIDPDKCTHLMYGFADIKYKTIQVDPYQDDNHN 92
DB 23 KLVCYTSWSQYREGVSGFLPDIAIQPFLCTHIIYSFANIS-----SDN 70
QY 93 E---KRGYERFNNRLKNPELTMTISLGWYEGSKYSDMAANPTYRQFIQSVD 149
DB 71 EWNDSNYDKLNKLKTRNTNLTKLLSVGQWKGKFRFSEIASNTERRTAFVRSV 130

| | | | |
|-----|----|--|-----|
| 150 | QY | XYFEDGLDLDWEYFGSRLGNPKDKQNYALVRELKDAF-----EP--HGYYLLTAASVFGKD | 203 |
| 131 | Db | YGFDEGLDLAWLFRUR-----DKOYFTLLIKELNAEFTKEVQFGREKLLLSAALSAGKV | 184 |
| 204 | QY | KIDRAYDIKELNKLFDPMNVMTDHYHGWFENFYGHNAFLVKNRDEIDELHTYENNVNYTMH | 263 |
| 185 | Db | AIDTGYDIAIAQHLDFINLMTYDEHGVWRQITGHHSFLPQQQKDT--RPRDSNVNYAVQ | 243 |
| 264 | QY | YYLNGATRDKLVMGVPPFYGRAMWSTEDRSKLUKDPAGKMSPGQFTSGEGVLSUYTELCO | 323 |
| 244 | Db | YMTRLGAQASKLLNGIPTFGKSFTLAS--SENQLGAPISGEGFLPGRFTKEAGTLAYVEICD | 302 |
| 324 | QY | LFQKEBWHIQYDEYNNAPYGVNDKLTWGVGYDLASICKLAFLKELGVSGV-----MWSLSE | 379 |
| 303 | Db | FLKGAEBVHLRSHN--KVPFATKGNQWGYEDKESVKYKVGFLKEKCLAGAGCWGHWMI | 359 |
| 380 | QY | NDDFKGHCGPKNPLLNKVNHMINGDEKNSFCEILGPSTPTPTPTPT | 426 |
| 360 | Db | ---SGHCQPKN-----SSRPTPTSRMP | 378 |

RESULT 12

S27879

secretory protein YM-1 precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999

C/Accession: S27879

R/Chang, N.C.A.; Liu, C.H.; Chang, A.C.

submitted to the EMBL Data Library, June 1992

A/Description: Molecular characterization of a secretory protein (YM-1) transiently expressed in

A/Reference number: S27879

A/Accession: S27879

A/Molecule type: mRNA

A/Residues: 1-399 <CHA>

A/Cross-references: EMBL:M94584; NID:G202441; PID:AA862394.1; PID:G202442

C/Superfamily: Streptomyces chitinase chi40

F/1-21/Domain: signal sequence #status predicted <SIG>

F/22-399/Product: secretory protein YM-1 #status predicted <MAT>

[illegible]

RESULT 13

A38221
chitinase (EC 3.2.1.14) MFL - nematode (Brugia malayi)
C:Species: Brugia malayi
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A38221
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Piessens, W.F.; Perler, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmucosal-blocking antibodies recognize microfilarial chitinase
A:Reference number: A38221; MUID:92179220; PMID:1542646
A:Accession: A38221
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <PU>
A:Cross-references: GB:M73689; NID:gl56063; PIDN:AAA27854.1; PID:gi56064
A:Note: sequence extracted from NCBI backbone (NCBIP:95345)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

| | | |
|----|---|---|
| | Query Match | 19.5%; Score 604.5; DB 2; Length 504; |
| | Best Local Similarity | 30.1%; Pred. No. 3e-30; |
| | Matches | 156; Conservative 86; Mismatches 208; Indels 69; Gaps 17; |
| Qy | 37 CYVGTWSVIYKVD - PYTIEDIDPFCKTHLMYGFAKDXYKYTTQVPDYO-DNNHNSWEK 94 : : : : : : : | |
| Dd | 27 CYYTNQAQRDGEGFLPGNIPNGLCTHILYAKUDE---LGDSKPFEWNDEDTEWSK 82 : : : : : : : | |
| Qy | 95 RGYERFNNLBKNPELTMTISLGWZGSGSKSDMAANPYRROFTQSULDFIQEYKFDG 154 : : : : : : : | |
| Dd | 83 GWYSAVTKLRETNPGLKVLLSYGYNPFGSAIFGTIAKSQAOKTFRIKSAIFAIRKKNFDG 142 : : : : : : : | |
| Qy | 155 LLDLWEYP-CSRIGNPKDKQNVALVRELKDAPFPHGILLTAASPGKDIIDRAYDIKE 213 : : : : : : : | |
| Dd | 143 FDLDEWPVGVBEBHAKLVEAMKTAFBEBAKTSGKOR-LLLTAARVSAGKGITDGSYNVES 201 : : : : : : : | |
| Qy | 214 LNKLFDMMVMVTDYDHGGWFYGHNAPLYKRPDDETDELTYFNVNVTMYVLNNGATRD 273 : : : : : : : | |
| Dd | 202 LGKNFDLLFLVSLDHLGSEWKNDVLHGKHPTKGEVSGI-GIENTEAAADYASKGMPPKE 260 : : : : : : : | |
| Qy | 274 KLVMGVFPYGRANSIEBRSKLKGDPAKMGSPPGFTSCEBGVLASYIELCOLFOKEEWHIQ 333 : : : : : : : | |
| Dd | 261 KIIGIPMVAQGWLTDNPSETA.GAAARSRSASKTNPAGGTASYEICKILEGGGETV 320 : : : : : : : | |
| Qy | 334 YDEYNAPYGVNDKIWYGVDLASICKLAFLKELGVSVGWMSLENDDFKG-HCGP-KN 391 : : : : : : : | |
| Dd | 321 HQEGVGA-YMWGDOWGYNEETIRIKMKWLKEKGYGGAFIWALDDFDTGSKSCGKGPY 379 : : : : : : : | |
| Qy | 392 PLLNKVHNMDGBKSNFECILGPSITTPPTTPPTTPPTTPPTTPPTTPPTTPPTTPPT 451 : : : : : : : | |
| Dd | 380 PLLNAISSLEGGSEN-----PEITTEPSITEAYETDEEE----- 418 : : : : : : : | |
| Qy | 452 TTPSFPTTPPTTPPTTPPTTPAPTSTPSPSTPTTHTSETPKYTIV----VDGHLIKC 507 : : : : : : : | |
| Dd | 419 -----TSETEAYDTD-----ETEETSET-EATTYDDETBGO--EC 451 : : : : : : : | |
| Qy | 508 -YKEGDIPHNTIHXYLCBFVGNWWHVIMPCPPGIW 545 : : : : : : : : | |
| Dd | 452 PERGLEPHEDCHLFLOC-----ANNIAYMCOPATFFF 486 : : : : : : : : | |

RESULT 14
DB3764
chitinase BH0916 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: DB3764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: AB3650; MUID:20512582; PMID:11059132

[illegible]

Search completed: March 22, 2004, 07:01:30
Job time : 38.0951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:31:13 ; Search time 21.2868 Seconds
(without alignments)
1357.597 Million cell updates/sec

Title: US-09-662-293-18

Perfect score: 3107

Sequence: 1 MKTIYALISIMACIGLNAS.....IMCPPGTIWQCKLTCIGE 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 838.5 | 27.0 | 554 | 1 | CHIT MANSE |
| 2 | 738 | 23.8 | 466 | 1 | CHT1 HUMAN |
| 3 | 734.5 | 23.6 | 473 | 1 | CHIA MOUSE |
| 4 | 723 | 23.3 | 508 | 1 | CHIL DROME |
| 5 | 708.5 | 22.8 | 476 | 1 | Q9W5U3 drosophila |
| 6 | 705.5 | 22.7 | 617 | 1 | Q9B9P6 homo sapien |
| 7 | 696.5 | 22.4 | 527 | 1 | Q11174 caenorhabdi |
| 8 | 685.5 | 22.1 | 537 | 1 | Q28990 sus scrofa |
| 9 | 685 | 22.0 | 539 | 1 | Q28042 bos taurus |
| 10 | 680 | 21.9 | 381 | 1 | Q28542 ovis aries |
| 11 | 680 | 21.9 | 671 | 1 | Q61362 mus musculus |
| 12 | 680 | 21.9 | 721 | 1 | Q60557 mesocricetu |
| 13 | 668 | 21.5 | 678 | 1 | Q62010 mus musculus |
| 14 | 662 | 21.3 | 383 | 1 | Q12889 homo sapien |
| 15 | 654.5 | 21.1 | 623 | 1 | P36222 homo sapien |
| 16 | 630 | 20.3 | 390 | 1 | P36718 papio anubi |
| 17 | 611.5 | 19.7 | 396 | 1 | Q15782 homo sapien |
| 18 | 604.5 | 19.5 | 504 | 1 | Q9d7q1 mus musculus |
| 19 | 576 | 18.5 | 458 | 1 | P29030 brugia mala |
| 20 | 561 | 18.1 | 699 | 1 | Q9W5U2 drosophila |
| 21 | 411 | 13.2 | 423 | 1 | P20323 bacillus ci |
| 22 | 410.5 | 13.2 | 423 | 1 | P48827 trichoderna |
| 23 | 407.5 | 13.1 | 563 | 1 | P32470 aphanoclad |
| 24 | 384 | 12.4 | 820 | 1 | P07254 serratia ma |
| 25 | 377.5 | 12.1 | 550 | 1 | P32823 alteromonas |
| 26 | 360.5 | 11.6 | 551 | 1 | Q10363 orgyia pseu |
| 27 | 353 | 11.4 | 499 | 1 | P41684 autographa |
| 28 | 341 | 11.0 | 619 | 1 | P11797 serratia ma |
| 29 | 339.5 | 10.9 | 427 | 1 | P36909 streptomyce |
| 30 | 326 | 10.5 | 610 | 1 | P54196 coccidioid |
| 31 | 280 | 9.0 | 1046 | 1 | P11220 streptomyce |
| 32 | 266.5 | 8.6 | 474 | 1 | P96156 vibrio furn |
| 33 | 259 | 8.3 | 5179 | 1 | P19275 thermoprote |
| | | | | | Q02817 homo sapien |

ALIGNMENTS

RESULT 1

CHIT MANSE

| ID | CHIT MANSE | STANDARD | PRT | 554 AA. |
|----|--|----------|-----|---------|
| DT | 01-JUN-1994 (Rel. 29, Created) | | | |
| DT | 01-JUN-1994 (Rel. 29, Last sequence update) | | | |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update) | | | |
| DE | Endochitinase precursor (EC 3.2.1.14). | | | |
| OS | Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea; | | | |
| OC | Sphingidae; Sphinginae; Manduca. | | | |
| OX | NCBI_TaxID=7130; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RP | MEDLINE=93357793; PubMed=8353525; | | | |
| RA | Kramer K.J., Choi H.K., Choi H.K., Muthukrishnan S.; | | | |
| RT | "Sequence of a cDNA and expression of the gene encoding epidermal and | | | |
| RT | gut chitinases of Manduca sexta."; | | | |
| RL | Insect Biochem. Mol. Biol. 23:691-701(1993). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RP | MEDLINE=97215580; PubMed=9061927; | | | |
| RA | Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.; | | | |
| RT | "Isolation and characterization of a genomic clone for the gene of an | | | |
| RT | insect molting enzyme, chitinase."; | | | |
| RL | Insect Biochem. Mol. Biol. 27:37-47(1997). | | | |
| CC | !- FUNCTION: Digest chitin in the exoskeleton during the molting | | | |
| CC | process. | | | |
| CC | !- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N- | | | |
| CC | acetyl-D-glucosamine polymers of chitin. | | | |
| CC | !- SUBCELLULAR LOCATION: Secreted. | | | |
| CC | !- TISSUE SPECIFICITY: Epidermis and gut. | | | |
| CC | !- DEVELOPMENTAL STAGE: High levels seen in the epidermis on day 0, | | | |
| CC | but rapidly disappears and is undetected on days 1-4 of fifth | | | |
| CC | instar. It reappears on day 5 and peaks on day 7 after which a | | | |
| CC | rapid decline is seen. In the gut is detected on day 6 with lower | | | |
| CC | levels seen on days 0, 7 and 8. | | | |
| CC | !- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl | | | |
| CC | hydrolases). | | | |
| CC | !- SIMILARITY: Contains 1 chitin-binding type-2 domain. | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.ies-sib.ch/announce/ | | | |
| CC | or send an email to license@sib-sib.ch). | | | |
| CC | ----- | | | |
| CC | EMBL; U02270; AAC04924.1; - | | | |
| DR | EMBL; I49234; AAB53952.1; - | | | |
| DR | PIR; A56596; A56596. | | | |
| DR | InterPro; IPR002557; Chitin bind_Pera. | | | |
| DR | InterPro; IPR001223; Glyco_hydro_18. | | | |
| DR | InterPro; IPR001579; Glyco_hydro_18AS. | | | |

Q62635 rattus norv
Q05049 xenopus lae
O10341 orgyia pseu
P10667 xenopus lae
P22533 caldocellum
P14918 zea mays (m
Q06885 dictyosteli
P47179 saccharomyc
P22699 dictyosteli
P09805 kluyveromyc
P24152 sorghum bic
P02840 drosophila

```
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitED2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00940; CHIT_BIND_II; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 554 ENDOCHITINASE.
FT DOMAIN 396 453 SER/THR-RICH.
FT DOMAIN 495 553 CHITIN-BINDING TYPE-2.
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 554 AA; 62203 MW; 3989D756C96CD490 CRC64;

Query Match 27.0%; Score 838.5; DB 1; Length 554;
Best Local Similarity 35.4%; Pred. No. 6.3e-39;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;

QY 6 ALLSTMAGIGLMMNASIKRDHNDYSKNPMRIYCVYGVTSVYHK-VDPYTIEDIDPFKCTHL 64
DQ 3 ATLALVALATAV-----QDSRARIVCYFSNMAVYRPGVGRYGIEDIPVEKCTHI 55
QY 65 MYGFAKIDYKYTIQVDPYQDDNHNWKKRGYERFNNLRKNPELMTTMSLGGWYEGSE 124
DQ 56 IYSGIVTEGSEVLIIIDPELD-----VDKNGFENFTSLRSHSPVSKFMVAVGWAAGSS 110
QY 125 KYSDMAANPTVROQFIQSVLDFLOEYKPDGLDLDWEYPSGR-LGNPKIDKQNYALVREL 183
DQ 111 KYSHVAQKSTRMSFIRSWFLKKYDPDGLDLDWEYFGAARDGSPSFKDKFLYLVOEL 170
QY 184 KDAP--EHBGYLLTAASVPGKDKIDRAYDIKELNKLDPMMWNTYDVHGGWENFYGNAP 241
DQ 171 RRAFTRVGKGWELTAAPLANFRLMGEGYHVPELCOELDAIHVMYDLRGNWAGFADVHSP 230
QY 242 LYKRDETEDEHTYFNNVNTWHYLLNNGATDKLVGMVPPFYGRAMSTED----- 290
DQ 231 LYKRPHD-QWAYEKLNVNDGLHLWEEKCPNKLVLGIPFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLGLDPAKGMSPPGISGEEGLSYIELCOLFKQEE--WHIQDYEVNAPYGYNDKI 348
DQ 290 NKEAGGDPAPYATNATG-----WAYEICTEVDKDDSGTWKWDQKQKCPYAKGTQ 342
QY 349 WGYDDLASISCKLAFKLKELGVGMVWNSLENDFKHCGKPNPLNKNVNMINGDEKNS 408
DQ 343 WGYEDPRSVEIKMNIQKQYLGANTWALDMDDFQGLCGEKPLIKLHKRMS----- 396
QY 409 FECILGPSTTPTPTTPTTPTT-----TTPTTPS-----PTTPTTPSTTTTPTSP 456
DQ 397 -----SYTVPPPHPTNTTTPPEWARPSPSPSEGDPIPTTTAKPASTTKTTVKTT 449
QY 457 TPTTTPSPFT-----PTP-TTPTPAPTSTPSPTTTEHTSETPKYTTTVVDGHLI 505
DQ 450 TTTAKPQSVIDEENDINVRPEKPEQPPEVEVP-PTENE-----VDGSEI 497
QY 506 KYCEGD-IPHTNTHKYLVCVFVNGWGWVHIMFCPPGTIWCQEKLTIC 552
DQ 498 -CNSQDDYIPDKKHCDKWRC--VNGE--AMQFSQCHQTVFVNLNVC 540

RESULT 2
ID CHIT1 HUMAN
AC Q13231; O9H3V8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chitotriosidase 1 precursor (EC 3.2.1.14) (Chitinase 1).
```

```
GN CHIT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND SUBCELLULAR
RP LOCATION.
RC TISSUE=Macrophage;
RX MEDLINE=96064695; PubMed=7592832;
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J.,
RA Aerts J.M.F.G.;
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT produced by macrophages.";
RL J. Biol. Chem. 270:26252-26256 (1995).
RN [2]
RP SEQUENCE OF 22-42 AND 178-198, FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=95138187; PubMed=7836450;
RA Renkema G.H., Boot R.G., Muijsers A.O., Donker-Koopman W.E.,
RA Aerts J.M.F.G.;
RT "Purification and characterization of human chitotriosidase, a novel
RT member of the chitinase family of proteins.";
RL J. Biol. Chem. 270:2198-2202 (1995).
RN [3]
RP POLYMORPHISM, AND DESCRIPTION OF ISOFORM 3.
RX MEDLINE=98421482; PubMed=9748235;
RA Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliet J.,
RA de Meulener T.M.A.M.O., Mannens M.M.A.M., Aerts J.M.F.G.;
RT "The human chitotriosidase gene. Nature of inherited enzyme
RT deficiency.";
RL J. Biol. Chem. 273:25680-25685 (1998).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 22-386 IN COMPLEX WITH
RP CHITIOSE AND ALLOSAMIDIN.
RX MEDLINE=22095530; PubMed=11960986;
RA Fusetti F., von Moeller H., Houston D., Rozeboom H.J., Dijkstra B.W.,
RA Boot R.G., Aerts J.M.F.G., van Aalten D.M.F.;
RT "Structure of human chitotriosidase. Implications for specific
RT inhibitor design and function of mammalian chitinase-like lectins.";
RL J. Biol. Chem. 277:25537-25544 (2002).
CC -I- FUNCTION: Degrades chitin and chitotriose. May participate in the
CC defense against nematodes and other pathogens. Isoform 3 has no
CC enzymatic activity.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -I- SUBUNIT: Monomer.
CC -I- SUBCELLULAR LOCATION: Secreted. A small proportion is lysosomal.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q13231-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q13231-2; Sequence=VSP_008631, VSP_008632;
CC Name=3;
CC IsoId=Q13231-3; Sequence=VSP_008633;
CC Note=Duplication of 24 bp in exon 10 leads to the use of a
CC cryptic splice site. The normal splice site is still present but
CC not used.
CC -I- TISSUE SPECIFICITY: Detected in spleen. Secreted by cultured
CC macrophages.
CC -I- POLYMORPHISM: A 24 bp duplication in exon 10 leads to the
CC activation of an alternative splice site and the production of an
CC inactive protein. About 6% of the population are deficient for
CC CHIT1 activity, while 35% are carriers and show reduced enzyme
CC levels. PEOPLE with CHIT1 deficiency appear perfectly healthy.
CC -I- MISCELLANEOUS: Patients with type 1 Gaucher disease (GD 1)
CC [MIM:230800] have very high plasma levels of CHIT1, and this can
CC be used as diagnostic aid and to evaluate the success of
CC treatment. Successful therapy brings the CHIT1 activity levels
CC back to normal.
CC -I- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -I- SIMILARITY: Contains 1 chitin-binding type-2 domain.
```

| | | | |
|----|---|---|-------------------------|
| | ----- | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | ----- |
| CC | DR | EMBL; U29615; AAC50246.1; ; | |
| CC | DR | EMBL; U62862; AAG10644.1; ; | |
| CC | DR | PDB; 1LGUV; 31-JAN-02. | |
| CC | DR | PDB; 1ILG1; 18-SEP-02. | |
| CC | DR | PDB; 1ILG2; 18-SEP-02. | |
| CC | DR | PDB; 1LLQ0; 29-JUL-03. | |
| CC | DR | Genew; HGNC:1936; CHIT1. | |
| CC | DR | MIM; 600031; ; | |
| CC | DR | GO; GO:0005615; C:extracellular space; TAS. | |
| CC | DR | GO; GO:0004568; F:chitinase activity; TAS. | |
| CC | DR | GO; GO:0009617; P:response to bacteria; TAS. | |
| CC | DR | GO; GO:0009613; P:response to pest/pathogen/parasite; TAS. | |
| CC | DR | InterPro; IPR002557; Chitin_bind_PeraA. | |
| CC | DR | InterPro; IPR001223; Glyco_hydro_18. | |
| CC | DR | InterPro; IPR001579; Glyco_hydro_18AS. | |
| CC | DR | Pfam; PF01607; CBM 14; 1. | |
| CC | DR | Pfam; PF00704; Glyco_hydro_18; 1. | |
| CC | DR | ProDom; PD000471; Glyco_hydro_18; 1. | |
| CC | DR | SMART; SM00494; ChtBD2; 1. | |
| CC | DR | SMART; SM00636; Glyco 18; 1. | |
| CC | DR | PROSITE; PS00940; CHIT BIND II; 1. | |
| CC | DR | PROSITE; PS01095; CHITINASE_18; 1. | |
| KW | Carbohydrate metabolism; Chitin degradation; | | |
| KW | Polysaccharide degradation; Hydrolase; Glycosidase; Chitin-binding; | | |
| KW | Signal; Alternative splicing; 3D-structure. | | |
| FT | SIGNAL | 1 21 | CHITOTRIOSIDASE 1. |
| FT | CHAIN | 22 466 | CHITIN-BINDING TYPE-2. |
| FT | DOMAIN | 417 466 | BY SIMILARITY. |
| FT | ACT SITE | 140 140 | |
| FT | DISULFID | 26 51 | |
| FT | DISULFID | 307 370 | |
| FT | VARSPLIC | 386 387 | SL-> NG (in isoform 2). |
| FT | VARSPLIC | 388 466 | /FTid=VSP_008631. |
| FT | VARSPLIC | 344 372 | Missing (in isoform 2). |
| FT | VARSPLIC | | /FTid=VSP_008632. |
| FT | VARSPLIC | | Missing (in isoform 3). |
| FT | VARSPLIC | | /FTid=VSP_008633. |
| FT | SEQUENCE | 466 AA; 51681 MW; B4312DI_E885E386D CRC64; | |
| | Query Match | 23.8%; Score 738; DB 1; Length 466; | |
| | Best Local Similarity | 34.2%; Pred. No. 1.6e+33; | |
| | Matches 152; Conservative | 85; Mismatches 150; Indels 58; Gaps 11; | |
| Qy | 34 RIVCVGVSWYHKVDP-YTTIEDDPKCTHMYGFAKIDEYKYTIQVFDPYQDDNENSW 92 | | |
| Db | 23 KLVCFYNWAQRGEARFLPKDLDPSCITHLIYAFAGMTNHQLST-----TEW 71 | | |
| Qy | 93 -EKGYERFNRLRKNPELTMI SLGGWYEGSEKYSMDAANPTVRQFIOSVLDFLOEYK 151 | | |
| Db | 72 NDETLYQEFGNLKKXNPKLTLIAIGWNFGTKQFTDMVAANNRQTFSNAIRFLRKYS 131 | | |
| Qy | 152 FDGLLDLWEYSGSRGNPKDKQNLYALVRELKDAPEPHG-----YLITAAVSPEGDK 204 | | |
| Db | 132 FDGLLDLWEYPSGS-GSPAADKERFTTLTVQPLANAFOOEQAQTSGERILLISAAPAGQTY 190 | | |
| Qy | 205 IDRAYDIKELNKLPDMNMVMYDYVHGWNFYGHNAPLYKRPFDTDELHTVENVNVTMHY 264 | | |
| Db | 191 VDAGEYVDKIQNLDLFVNLMAIDFHGSWEKVTHNSPLYLKQEESSGAAS-LNVDAAVQQ 249 | | |
| Qy | 265 YLNNGATRDKLVMGVPPFYGRAWSIETRDSKLGLGPAPKMGSPPGPISGEVLSYIELCQL 324 | | |
| Db | 250 WLQGTTPASKILLGNPTYGRSFTTLASSSDITEVGAPATGSGTFGPTKEGMLAYEVCWS 309 | | |
| Qy | 325 FQKEWHIOXDYEYNAPYGYNDKIWLWGYYDIIASICKLAFLKELGVSGVMVMSLENDFFK 384 | | |

[illegible]

```
DR PROSITE; PS00940; CHIT_BIND II; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Carbohydrate metabolism; Chitin degradation;
KW Polysaccharide degradation; Hydrolase; Glycosidase; Chitin-binding;
KW Signal; Alternative splicing.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 476 ACIDIC MAMMALIAN CHITINASE.
FT DOMAIN 427 476 CHITIN-BINDING TYPE-2.
FT DOMAIN 415 420 POLY-SER.
FT ACT_SITE 140 140 BY SIMILARITY.
FT DISULFID 26 51 BY SIMILARITY.
FT VARSPLIT 307 372 BY SIMILARITY.
FT VARSPLIT 1 161 Missing (in isoform 3).
FT VARSPLIT 1 108 Missing (in isoform 2).
FT VARSPLIT 339 339 I -> V (IN REF. 1).
FT CONFLICT 432 432 V -> G (IN REF. 1).
SQ SEQUENCE 476 AA; 52271 MW; 92B27BAD2F7EB4CC CRC64;

Query Match 22.8%; Score 708.5; DB 1; Length 476;
Best Local Similarity 35.9%; Pred. No. 6.4e-32;
Matches 149; Conservative 75; Mismatches 154; Indels 37; Gaps 10;

QY 34 RIVCVGTVSVYHK-VDPYTTIEDIDPKCTHLMYGEAKIDYKYTIQVDFPYQDDNHSW 92
DB 23 QLTCTFTNWAQYRGLGRFMDNIDPCLCTHLIVAFAGRQNNETT-----IEW 71

QY 93 -EKGYERFNNLRKNPELTMTISLGWYEGSEKYSDMAANPTYRQFIQSVLDFLOEYK 151
DB 72 NDVTLYQAFNGKKNKNSQLTKLLAIGGNWFTAFATMVSTPENRQTFITSVIKFLQYE 131

QY 152 FDGLDWEYPGSLGNPKIDQNYLALVRELKDAFPHG-----YLTAASVSPGKDX 204
DB 132 FDGLDWEYPGSR-GSPDQKHLFTVLVQSMREAFQEAQINKPLMTAAVAAGISN 190

QY 205 IDRAYDIKELNKLFDWMNVTYDHGGWENFYGHNAPLYKRPDSTDLHYFVNVYTHY 264
DB 191 IQSGYEIPQLSQYLDYHWTYDLHGGWEGTGENSPLYKPTDTS-NAYLNVYVYNNY 249

QY 265 YLNGATKDLVMGVFFYGRAWSIEDRSKLGIDPAKGMSPPGFISGEEGLSVIELCOL 324
DB 250 WKGNGAPAEKLVIFPTTYGHNFILSNPNTGIGAPTSAGPAGYAKESGIWAYEIC-T 308

QY 325 FQKEBWHIYDEYNAPYGYNDKIWGVYDDILASISCKLAFKLKELGVSQVWVMSLENDFFK 384
DB 309 FLKNGATQGWADAPQVEVPAYQGVNVGYNDIKSFIDIKAOWLKHNKFGAMWALDLDFT 368

QY 385 G-HGCP-KNPLLNKVNHNMGDEKNSFECILGPTSTPTPTPTPTPTPTPTPTPTPT 437
DB 369 GTFCNQGRFPLI-----STLKALGLQASCTAPAQPIEPTAAPSGS 411
```

RESULT 6

```
ID CHIT_CABEL STANDARD; PRT; 617 AA.
AC Q11174; O17321;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable endochitinase (EC 3.2.1.14).
GN CHT-1 OR C04F6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nham M.;
RL Submitted (DSC-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 62-178 FROM N.A.
```

```
RX MEDLINE=98324849; PubMed=9662472;
RA de la Vega H., Specht C.A., Liu Y., Robbins P.W.;
RT "Chitinases are a multi-gene family in Aedes, Anopheles and
RT Drosophila.";
RL Insect Mol. Biol. 7:233-239 (1998).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -!- SIMILARITY: Contains 2 chitin-binding type-2 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; U42835; AA893586.1; -
DR EMBL; AF026152; AA881847.1; -
DR PIR; T15408; T15408.
DR PIR; T37249; T37249.
DR WormPep; C04F6.3; CE03923.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18A.
DR Pfam; PF01607; CEM_14; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtB2; 2.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00940; CHIT_BIND II; 2.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Repeat.
FT DOMAIN 478 534 CHITIN-BINDING TYPE-2 1.
FT DOMAIN 563 617 CHITIN-BINDING TYPE-2 2.
FT ACT_SITE 179 179 PROTON DONOR (BY SIMILARITY).
FT DOMAIN 435 478 THR-RICH.
FT CONFLICT 138 138 W -> C (IN REF. 2).
FT CONFLICT 163 165 ITF -> TTS (IN REF. 2).
FT CONFLICT 176 176 I -> L (IN REF. 2).
SQ SEQUENCE 617 AA; 66857 MW; DDALD2AAAC0ES4DA CRC64;

Query Match 22.7%; Score 705.5; DB 1; Length 617;
Best Local Similarity 30.2%; Pred. No. 1.3e-31;
Matches 178; Conservative 90; Mismatches 210; Indels 111; Gaps 19;

QY 37 CVYGVTVSVYHK-VDPYTTIEDIDPKCTHLMYGEAKIDYKYTIQVDFPYQDDNHSWEKR 95
DB 57 CYFTNWAQYRQGRKFPVSDYTPGLCTHLIPAFGWNA-DYTVRAYDP--ADLPNDWAGE 113

QY 96 G-YERNRLRNKPELTMTISLGWYEGSEKYSDMAANPTYRQFIQSVLDFLOEYKPDG 154
DB 114 GMYRVNKLKVTDTQLTKLLSFGWSFGTALPQGAASASRKFVIDSAITVRTWTFDG 173

QY 155 LLDLWEYPGSLGNPKIDQNYLALVRELKDAFPHG-----YLTAASVSPGKDKIDR 207
DB 174 IDIDWEYPSG-----ATDMANYVALVKEKLAACEASEAGSTGKDRLLVTAAGAATIDA 228

QY 208 AYDIKELNKLFDWMNVTYDHGGWENFYGHNAPLYKRPDSTDLHYFVNVYTHYILN 267
DB 229 GVDIPNLAFNDFPILLMSYDFTGAWASLVGFNSPLYATTELPAENGW-NVDSSARYNQ 287

QY 268 NGATRKLVGMVFFYGRAWSIEDRSKLGIDPAKGMSPPGFISGEEGLSVIELCOLFOK 327
DB 288 KGMPEKILVGMPTYGRGWTNNASAINPGTSGSPAKITQYVQ-EAGVGAYFECEMLAN 346

QY 328 EWHIYQDEYNAPYGYNDKIWGVYDDILASISCKLAFKLKELGVSQVWVMSLENDFFKHC 387
DB 347 GATR-YWDSQSQVPIVQGNQWMSYDDEESFANKYAYKREGYGFVWTLDFDFFNAGC 405

QY 388 GPKN-----PFLNKNVHNMGDEKNSFECIL---GFSTTPTT-----PTTPTPTT 429
```



```
QY 188 EPHG-----YLLTAAVSPGKDIDRAYDIKEINLKLFDWMNVTYDHYGWFENFYGHNA 240
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 168 KNEAQLTMRPRILLSAVSDPHVIOKAYDARILGRLDLFISVSLDGLGSEKVTGHS 227
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 PLYKRPDDELTHYFNVTMYLYLNGNATRKLVGVFPFYGRAWSIEDRSKILGDPDA 300
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 228 PLFSLPGDPK-----SSAYAMSVYRQLGVPEKLLMGLPTYGRTHLLRASQNELGAGA 281
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 KMSPPQFISGEGVLSYIELCOLFQKEWHIQDYNNAPYNDKINWGYDDLASISC 360
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 282 AGPASPGKYTKOAGFLAYVYCVFORAKRWINDQY--VPYAFKGEWVGYDDAISFGY 339
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 KLAFLKELGVSGVWVSLNDDFKGH--CGP-KNPLLNKVNMMINGEKKSFECLGPSIT 418
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 340 KAFPIKREHFGGKAVWTLDDDDFRGNFCGTPPPLAHTLNNLLVNDEFFS----- 389
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 419 TPTPT-----TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 466
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 390 TSPKFWFSTAVNSRIGPEWPTWTRDLTTGLILPLGGEAVATETHRSATWTTTPRGE 449
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 467 TPTPT-TP-----TPAPTSTPTPTTHTS 491
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 450 TATPTRTPLSSGRRTAAPEGKTSPGKPLTS 481
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
C3L1 MOUSE
ID _C3L1_MOUSE STANDARD; PRT; 381 AA.
AC Q61362:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE (GP-39) (BRP39 protein).
GN CHI3L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic markers
RT generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -1- FUNCTION: May play an important role in the capacity of cells to
CC respond to and cope with changes in their environment (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -1- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X93035; CAA63603.1; --
CC PIR: S61551; S61551.
CC MGD; MGI:1340899; Chi3l1.
CC InterPro; IPR001223; Glyco_hydro_18.
CC InterPro; IPR001579; Glyco_hydro_18AS.
CC Pfam; PF00704; Glyco_hydro_18_1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC SMART; SM00636; Glyco_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
CC Glycoprotein; Signal.
CC SIGNAL 1 21
FT SIGNAL 1 21
```

```
FT CHAIN 22 381 CHITINASE-3 LIKE PROTEIN 1.
FT CARBOHYD 60 60 N-LINKED (GLCNAC..) (POTENTIAL).
SQ SEQUENCE 381 AA; 43001 MW; EF6581E184F0450 CRC64;

Query Match 21.9%; Score 680; DB 1; Length 381;
Best Local Similarity 37.1%; Pred. No. 1.8e-30;
Matches 141; Conservative 75; Mismatches 130; Indels 34; Gaps 10;

QY 34 RIVCYVGTWSVYHK-VDPYTTIEDIDPPKCTHLMYGFADIKYKTIQVDFPYQDDNHSW 92
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 KLVCFYTSWSQYREGVGSFLPDALQFLCTHIISFANIS-----SDNMLSTW 70
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 93 E---KRGYERFNNLRKLNPELTMTISLGGWYEGSEKYSDMAANPTYRQOFTQSVDLFLQE 149
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 EWNDESNDYKLNKLTNTNLTLLSVGVGWFGEKGFSEIASNTERRTAFVRSVAFPLRS 130
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 150 YKFDGLDLWEYPSGRLGNPKIDKONYLALVRELKDAF-----EP--HGYYLTAAVSPGKD 203
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 YGFDGLDLAWLYPLRL-----DKQYFSTLIKELNAEPTKEVQPGREKLLSAALSAGKV 184
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 204 KIDRAYDIKELNKLDPWNVMTYDHYGWFENFYGHNAPLYKRPDDELTHYFNVTMYTMH 263
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 AIDTGYDIAQAHLDFINLMTYDFHGVWRQITGHSPLFQCKDT-RFDYSYNNVAVQ 243
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 264 YYLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPKAGMSPGPFISGEGVLSYIELCQ 323
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 YMIRLGAQASLLMGIPTFGKSFTLAS--SENQLGAPISGEGLPGRFTKEAGTLAYEICD 302
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 324 LFOKEEHIQDYNNAPYNDKINWGYDDLASISCKLAFELKELGVSGVWVSLNDDDF 383
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 FLKGAEVHLSNE--KVFPATKGNQWGVYEHKESVKNKVGFLKEKLAGAMVWALDLDFF 360
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 384 KGHCGPKN--PILNKVHNNI 401
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 QGTCPQKEFFPLTNAIKDAL 380
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
OGP MESAU
ID _OGP_MESAU STANDARD; PRT; 671 AA.
AC Q60557; Q60526;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (3strogen-dependent oviduct protein) (ZP-0).
GN OVG1 OR OGP.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=96115007; PubMed=7492686;
RA Suzuki K., Sendai Y., Onuma T., Hoshi H., Hiroi M., Araki Y.;
RT "Molecular characterization of a hamster oviduct-specific
RT glycoprotein.";
RL Biol. Reprod. 53:345-354(1995).
RN [2]
RP SEQUENCE OF 22-671 FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=96192955; PubMed=8607967;
RA Paquette Y., Merlen Y., Malette B., Bleau G.;
RT "Allelic polymorphism in the hamster oviductin gene is due to a
RT variable number of mucin-like tandem repeats.";
RL Mol. Reprod. Dev. 42:388-396(1995).
RN [3]
RP SEQUENCE OF 14-671 FROM N.A., AND REVISIONS.
RC TISSUE=Oviduct;
RA Paquette Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
```



```
DR SMART; SMO0636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 721 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT DOMAIN 486 632 21 X 7 AA TANDEM REPEATS OF S-K-T-T-
[TAPE]-G-[IV].
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;

Query Match 21.9%; Score 680; DB 1; Length 721;
Best Local Similarity 33.1%; Pred. No. 3.6e-30;
Matches 175; Conservative 79; Mismatches 184; Indels 82; Gaps 16;

QY 34 RIVCVYGVTVSVVHKVDPIYI--EDIDPFKTHLMVGFADIKVYKTIQVDFPDQDNHNS 91
DB 23 KLVCFYTNWA-HSRPGASIMHDLDPFLCTHLIFAFASMSNNQI---VAKLQDENVL- 77
QY 92 WEKRGYERFNNLRKLNKPELTWMISLGGWYEGSEKYSMAANPTYRQOFTQSVLDLQBYK 151
DB 78 ----YPEFNKLNKERNRELKTLISIGWNFGTSRFTALSTLANREKFDISVIFLRING 132
QY 152 FDGLDLWDYPSGRIGNPKIDQNYLALVRELKDAPPHG-----YLLTAASVPGDK 204
DB 133 FDGLDLFFLYPGLR-GSPPHDRWNFLLEELQFAFEREALTOHPRLLSAAVSGIPSI 191
QY 205 IDRAYDIKELNKLFDWMVMTVDYHGGWENFYGHNAFLYKRPDETDELHTYFNVNVTWHY 264
DB 192 INTSYDALLGRLRDINVLSDHGSWEKFTGHNSPLFSLPESDK-----SSAYANNY 245
QY 265 YLNNGATRDKLVMGVFPFYGRAMSIDRSKLKGDPAKGMSPPGF-SGREGVLSYELCQL 324
DB 246 WRKLGTAPADKLINGFTYGRNFVLLKESKNGIQTASMGSPASPKYTKQAGFLAYEVCSF 305
QY 325 FQKEWHIYDYNNAPYNDKINWGYDDLASICKLAFLKELGVSGVMVMSLENDPDFK 384
DB 306 VQRAKKH--WIDYQYVYAFKGEMLGYDITISFSYKAMYKREHFGAGVWTLWDMDVR 363
QY 385 G-HCG-PKNPLNKVHNMINGENKSPFGLPSTTTPP----- 422
DB 364 GTFCGNGPPLVHILNELLVQTESN-----TLPQFWFTSSVNASGPGSENTAL 413
QY 423 ---TTPTPTPTPTPTPTPTTP-----TTTPS--PTPTPTPTPTPTPTPTPTPTPT 469
DB 414 TEVLITDTIKILPPGGEAMTTEVHRRYENMTVPDSGVTPCGTASPRKHAIVPENTMA 473
QY 470 PTPTPT-----APTSTPTPTPT---EHTSETPKYTYV 500
DB 474 AEAKTMTSLDFFSKITTYGSKTTTGISKITTYGSKITTYGV 513

ID OGP HUMAN STANDARD; PRT; 578 AA.
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Mucin 9).
GN OVGPI OR OGP OR MUC9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=95119256; PubMed=7819450;
RA Arias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
```

```
RT characterization of an estrogen-dependent human oviductal
glycoprotein.";
RL Biol. Reprod. 51:685-694(1994).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT GLN-676.
RA Jaffe R.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Coville G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
in the fertilization process and/or early embryonic development.
CC !- SUBCELLULAR LOCATION: Secretory granules.
CC !- TISSUE SPECIFICITY: Oviduct.
CC !- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; U09550; AAB04126.1; -
DR EMBL; U58010; AAB04126.1; -
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.
DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR EMBL; AL390195; CAC36039.1; -
DR MIM; 603578; -
GO; GO:0007565; P:pregnancy; TAS.
InterPro; IPR001223; Glyco_hydro_18.
InterPro; IPR001579; Glyco_hydro_18AS.
Pfam; PF00704; Glyco_hydro_18; 1.
ProDom; PD000471; Glyco_hydro_18; 1.
SMART; SMO0636; Glyco_18; 1.
PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Signal; Polymorphism.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 676 676 E -> Q (in dbSNP:7825).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;

Query Match 21.5%; Score 668; DB 1; Length 678;
Best Local Similarity 33.1%; Pred. No. 1.5e-23;
Matches 176; Conservative 91; Mismatches 192; Indels 72; Gaps 19;

QY 10 IMACIGLWNASIKEDHNDYKSNPMRIYCVYGVTVSVVHKVDPIYI--EDIDPFKTHLMYG 67
DB 4 LLLMWGLV---LVLLKHHDGAH--KLVCFYTNWA-HSRPGASIMHDLDPFLCTHLIFA 57
QY 68 FAKIDYKYTIQVDFPDQDNHNSWEKRGYERFNNLRKLNKPELTWMISLGGWYEGSEKYS 127
DB 58 PASMNNQI---VAKLQD-----EKLIPFNKLNKERNRELKTLISIGWNFGTSRFT 108
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:39:53 ; Search time 111.679 Seconds
(without alignments)
1568.003 Million cell updates/sec

Title: US-09-662-293-18
Perfect score: 3107
Sequence: 1 MKTIYAILGIMACIGLMNAS.....IMPCPGTIWCQKLTCTIGE 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL 25:*
- 1: sp_archea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|---------------------|
| 1 | 3107 | 100.0 | 555 | 5 Q9UGR7 | Q9UGR7 dermatophag |
| 2 | 1008.5 | 32.5 | 555 | 5 Q44079 | Q44079 anopheles g |
| 3 | 942 | 30.3 | 431 | 5 Q818H5 | Q818H5 araneus ven |
| 4 | 869 | 28.0 | 1635 | 5 Q17412 | Q17412 aedes aegypt |
| 5 | 854.5 | 27.5 | 4498 | 5 Q9W223 | Q9W223 drosophila |
| 6 | 847 | 27.3 | 544 | 5 Q9GQC4 | Q9GQC4 bombyx mori |
| 7 | 842 | 27.1 | 2838 | 5 Q8MP05 | Q8MP05 tenebrio mo |
| 8 | 839 | 27.0 | 566 | 5 Q8WR52 | Q8WR52 bombyx mori |
| 9 | 837.5 | 27.0 | 467 | 5 Q15993 | Q15993 penaeus jap |
| 10 | 837.5 | 27.0 | 543 | 5 Q9GR93 | Q9GR93 bombyx mori |
| 11 | 837.5 | 27.0 | 543 | 5 Q9GV05 | Q9GV05 bombyx mori |
| 12 | 837.5 | 27.0 | 565 | 5 P90710 | P90710 bombyx mori |
| 13 | 837.5 | 27.0 | 535 | 5 Q9VPR3 | Q9VPR3 drosophila |
| 14 | 832 | 26.8 | 460 | 5 Q9W2M7 | Q9W2M7 drosophila |
| 15 | 831.5 | 26.8 | 467 | 5 Q8ITU3 | Q8ITU3 penaeus van |
| 16 | 829.5 | 26.7 | 565 | 5 Q9GPG9 | Q9GPG9 bombyx mand |

| | | | | | |
|----|-------|------|------|-----------|---------------------|
| 17 | 823.5 | 26.5 | 557 | 5 Q8MTK0 | Q8MTK0 choristoneu |
| 18 | 823 | 26.5 | 488 | 13 Q90W34 | Q90W34 bufo japoni |
| 19 | 814.5 | 26.2 | 552 | 5 Q9GV44 | Q9GV44 spcdoptera |
| 20 | 810 | 26.1 | 574 | 5 Q17411 | Q17411 aedes aegypt |
| 21 | 802 | 25.8 | 483 | 5 Q23737 | Q23737 chelonius sp |
| 22 | 798 | 25.7 | 572 | 5 Q26042 | Q26042 penaeus jap |
| 23 | 790 | 25.4 | 620 | 5 Q9Y0D4 | Q9Y0D4 penaeus mon |
| 24 | 787 | 25.3 | 553 | 5 P91731 | P91731 hyphantria |
| 25 | 780 | 25.1 | 470 | 13 Q803B7 | Q803B7 brachydanio |
| 26 | 762.5 | 24.5 | 474 | 5 Q86L22 | Q86L22 lutzomyia l |
| 27 | 754 | 24.3 | 500 | 13 Q7ZV48 | Q7ZV48 brachydanio |
| 28 | 738 | 23.8 | 466 | 4 Q13231 | Q13231 homo sapien |
| 29 | 736.5 | 23.7 | 473 | 11 Q99PH2 | Q99PH2 mus musculu |
| 30 | 734.5 | 23.6 | 472 | 11 Q9JLN1 | Q9JLN1 mus musculu |
| 31 | 734.5 | 23.6 | 473 | 11 Q9D803 | Q9D803 mus musculu |
| 32 | 733 | 23.6 | 482 | 13 Q8AV87 | Q8AV87 gallus gall |
| 33 | 732.5 | 23.6 | 387 | 4 Q9H3V8 | Q9H3V8 homo sapien |
| 34 | 731.5 | 23.5 | 472 | 6 Q9SM17 | Q9SM17 bos taurus |
| 35 | 723.5 | 23.3 | 688 | 5 Q8WS85 | Q8WS85 drosophila |
| 36 | 722.5 | 23.3 | 1013 | 5 Q960M0 | Q960M0 drosophila |
| 37 | 717 | 23.1 | 527 | 5 P91773 | P91773 penaeus jap |
| 38 | 716 | 23.0 | 460 | 5 Q8WS95 | Q8WS95 glossina mo |
| 39 | 710 | 22.9 | 484 | 5 Q9W092 | Q9W092 drosophila |
| 40 | 708.5 | 22.8 | 476 | 4 Q9BZP6 | Q9BZP6 homo sapien |
| 41 | 703 | 22.6 | 929 | 5 Q8MY79 | Q8MY79 haemaphysal |
| 42 | 702.5 | 22.6 | 983 | 5 Q9VZV2 | Q9VZV2 drosophila |
| 43 | 691 | 22.2 | 462 | 5 Q9W2M6 | Q9W2M6 drosophila |
| 44 | 687 | 22.1 | 381 | 11 Q99J84 | Q99J84 mus musculu |
| 45 | 687 | 22.1 | 389 | 11 Q8BKL8 | Q8BKL8 mus musculu |

ALIGNMENTS

RESULT 1

Q9UGR7
ID Q9UGR7 PRELIMINARY; PRT; 555 AA.
AC Q9UGR7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE 98kDa HDM allergen.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcotiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Weber E.R., Hunter S., Steadman K., McCall C.;
RT "Cloning and Characterization of a 98 kDa Allergen from
RT Dermatophagoides farinae";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF178772; AAC52672.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PexA.
DR InterPro; IPR01223; Glyco hydro 18.
DR InterPro; IPR01579; Glyco hydro 18AS.
DR Pfam; PF00704; Glyco hydro_18; 1-
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 555 AA; 63238 MW; 0E4564A1A459B30B CRC64;

Query Match 100.0%; Score 3107; DB 5; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.9e-186;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKTIYAILSIACIAGLWNASIKRDHNDYSKNPMRIVCYVGTWSTVYHKVDYPTTIEDIDPFK 60
DB 1 MKTIYAILSIACIAGLWNASIKRDHNDYSKNPMRIVCYVGTWSTVYHKVDYPTTIEDIDPFK 60
QY 61 CTHLMYFAKIDYKTIQVDPDPYQDDNHNWSEKRGYERFNNLRLKNPELTMTISLGGWY 120
DB 61 CTHLMYFAKIDYKTIQVDPDPYQDDNHNWSEKRGYERFNNLRLKNPELTMTISLGGWY 120
QY 121 EGSKYSDMAANPYRQFIQSVLDLQEVKFDGLDLDWEYVPGSLGNPKIDKONYLALV 180
DB 121 EGSKYSDMAANPYRQFIQSVLDLQEVKFDGLDLDWEYVPGSLGNPKIDKONYLALV 180
QY 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNA 240
DB 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNA 240
QY 241 PLYKRPDETLHYFNNVMTYLLNGATRDKLVMGVPPFYGRAMSIEDRSKLKGDPA 300
DB 241 PLYKRPDETLHYFNNVMTYLLNGATRDKLVMGVPPFYGRAMSIEDRSKLKGDPA 300
QY 301 KGMSPPGFISGEGVLSYIELCOLFOKEEWHIOYDEVYNAPYGYNDKIWGYDDLASISC 360
DB 301 KGMSPPGFISGEGVLSYIELCOLFOKEEWHIOYDEVYNAPYGYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSMVMSLENDDFKGGCPKNPLNKHVNNMNGDEKNSFCILGPSTTTP 420
DB 361 KLAFLKELGVSMVMSLENDDFKGGCPKNPLNKHVNNMNGDEKNSFCILGPSTTTP 420
QY 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
DB 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
QY 481 TSPSTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCBFWNGGWVHIMPCP 540
DB 481 TSPSTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCBFWNGGWVHIMPCP 540
QY 541 PGTIWCQKLCIG 555
DB 541 PGTIWCQKLCIG 555

RESULT 2
O44079 PRELIMINARY; PRT; 525 AA.
ID O44079
AC O44079;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase.
GN AGCHI-1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gut;
RX MEDLINE=98030563; PubMed=9360958;
RA Shen Z, Jacobs-Lorena M.;
RT "Characterization of a novel gut-specific chitinase gene from the
RT human malaria vector Anopheles gambiae."
RL J. Biol. Chem. 272:28895-28900(1997).
DR EMBL; AF008575; AAB97764.1; -.
DR PIR; T4445; T4445.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008063; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolyase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
```

```
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 525 AA; 57211 MW; 3234360EBFF36165 CRC64;

Query Match 32.5%; Score 1008.5; DB 5; Length 525;
Best Local Similarity 38.4%; Pred. No. 4e-55;
Matches 211; Conservative 93; Mismatches 187; Indels 59; Gaps 13;

QY 8 LSIACIGLWNASIKRDHNDYSKNPMRIVCYVGTWSTVYHKVDYPTTIEDIDPFKCTHLM 65
DB 5 VGVVLVAVAAFAEPKGAASGKVKVCVGTWAVTPGNGRYDIEHIDPSLCTHLM 64
QY 66 YGFAKIDYKTIQVDPDPYQDDNHNWSEKRGYERFNNLRLKNPELTMTISLGGWEGSEK 125
DB 65 YGFFGNE-DATVRIIDPYLDLEN-WGRGHIKREVLKNGVPGGLKTLAAIGMNEGRK 122
QY 126 YSDMAANPYRQFIQSVLDLQEVKFDGLDLDWEYVPGSLGNPKIDKONYLALVRELKD 185
DB 123 FSMAASGELRKRFISDCVAFQCRGHGFDGLDWEYPAQRDGNPLIDRDNHAQLVEEMRE 182
QY 186 APEPHGYLLTAAVSPGKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKR 245
DB 183 EFDHYGLLLTAAVASVERPSAGSYDIPIKSKFHLNVMVYDMHGANDSYCGINAPLYRG 242
QY 246 PDDETDELHYFNNVMTYLLNGATRDKLVMGVPPFYGRAMSIEDRSKLKGDPAKGMSP 305
DB 243 SADTTDRLLQIINVNASIHFWLAQGGCTGRKVLGIPLYGRNFTLASAANTQIGAPTVGGT 302
QY 306 PGFISGEGVLSYIELCOLFOKEEWHIOYDEVYNAPYGYNDKIWGYDDLASISCLAF 365
DB 303 VGRYTRPGVMNGNEFEKLAATEAWDLWSEBQQVPYAVRNQWVGYDDLSVQLKVKYL 362
QY 366 KELGVSMVMSLENDDFKGGHC-GPKNPLNKHVNNMNGDEKNSFECILGPSTTPTTPT 424
DB 363 LQGLGGMVMSLETDPLGVCGGRYPLMHHSIRSLVNGT-----PSTTWPSPV 413
QY 425 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 484
DB 414 APTT-----STVAPGTTTTTPTGANPTQPTT--SDAPNHTTTTTEGNGPTTTPSG 466
QY 485 TTEHTSETPKYTYVDGHLIKCY--KEGDIHPHTNIHKYLVC-----EFVNGGWVH 535
DB 467 -----DG---FCAGGRYGFVPHPTNCARYICLTADTYVEFT----- 500
QY 536 IMPCPPTIW 545
DB 501 ---CPPTLF 507

RESULT 3
Q81SH5 PRELIMINARY; PRT; 431 AA.
ID Q81SH5
AC Q81SH5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase.
OS Araneus ventricosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Araneidae; Araneus.
OX NCBI_TaxID=182803;
RN [1]
RP SEQUENCE FROM N.A.
RA Han J.H., Kim S.R., Sohn H.D., Jin B.R.;
RT "Molecular cloning of a cDNA encoding the chitinase from the spider,
RT Araneus ventricosus."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV120879; AAN39100.1; -.
```

| | | | |
|--|--|----------|--|
| DR | GO; GO:0016787; F:hydrolase activity; IEA. | DR | GO; GO:0005576; C:extracellular; IEA. |
| DR | GO; GO:0005975; P:carbohydrate metabolism; IEA. | DR | GO; GO:0008061; F:chitin binding; IEA. |
| DR | InterPro; IPR001223; Glyco_hydro_18. | DR | GO; GO:000843; F:endochitinase activity; IEA. |
| DR | Pfam; PF00704; Glyco_hydro_18; 1. | DR | GO; GO:0016788; F:hydrolase activity, acting on glycosyl bonds; IEA. |
| DR | ProDom; PD000471; Glyco_hydro_18; 1. | DR | GO; GO:0005975; P:carbohydrate metabolism; IEA. |
| DR | SMART; SM00636; Glyco_18; 1. | DR | GO; GO:0006032; P:chitin catabolism; IEA. |
| DR | PROSITE; PS01095; CHITINASE_18; 1. | DR | InterPro; IPR002557; Chitin_bind_Pera. |
| SQ | SEQUENCE 431 AA; 47238 MW; 929439397B9BC23 CRC64; | DR | InterPro; IPR001223; Glyco_hydro_18. |
| Query Match 30.3%; Score 942; DB 5; Length 431; | | | |
| Best Local Similarity 43.8%; Pred. No. 4.6e-51; | | | |
| Matches 185; Conservative 78; Mismatches 135; Indels 26; Gaps 9; | | | |
| QY | 12 ACIGLM--NASIKRDHNDYKMPRIVCYVGTWVSVYHKVD-PYTIEDIDPFKCHLMYGF 68 | QY | 21 IKRDNDYKSNPM-----RIVCYVGTWVSVYHKVD-PYTIEDIDPFKCHLMYFAKIDE 73 |
| DB | 6 ACILLLVAVASQSRDRNQKKYKVCYLGSWANVRGGEKFLIEHIDPFKCHLMYGF 65 | DB | 650 INKETSNDENAIESDVYKVCYFTWAWYRQGNKGLPEDIDADLCTHIVYGFVLD 709 |
| QY | 69 AKIDEYKTIQVDPYODDNHNSWKRGRYERFNNLRKLNPELTMTSLGWNVEGSKYSD 128 | QY | 74 YKTIQVDPYQDDNNHNSW---EKRGYERFNNLRKLNPELTMTSLGWNVEGSKYSD 129 |
| DB | 66 AKLSNQ--IAVDPYLDLKEN-WGLGAFORFNNLRKLNPELTMTSLGWNVEGSKYSA 122 | DB | 710 DRLVTK-----PHDSWADIDNRFRYERVVEYKKGKKT--VAIGWNDSAGDKYSL 759 |
| QY | 129 MAANPYRQFIOGSLVDFLOEYKFDGLDWEYPSGLNPKIDKQNYLALVRELKDAFE 188 | QY | 130 RANPYRQFIOGSLVDFLOEYKFDGLDWEYPSGLNPKIDKQNYLALVRELKDAFE 185 |
| DB | 123 MAADPNARATFKVSDVDFCLKYDFDGLDMDWEYPANR-GGAADHKQNFVTLKELKEAFA 181 | DB | 760 VSAARAKQFADVVAFIEKYGFDGLDWEYPCWQVDCKKGFSDKEGFSALVVELSQ 819 |
| QY | 189 PHGVLLTAAPSPGKDKIDRAYDIKELNKLFDMMNVTYDVGWENFYGHNAPLYKRPDE 248 | QY | 186 AFEPHGILLTAAPSPGKDKIDRAYDIKELNKLFDMMNVTYDVGWENFYGHNAPLYK 245 |
| DB | 182 PHGLLSAAVSAAGNTITDIPGVAKYLDFFINWAYDLGSEWKTAGHNPALYERPE 241 | DB | 820 AFKPKGLLSAAVSPSKVDEGYDVTLSYMDMIAVAYDYGQWKKTGHVAPMVEH 879 |
| QY | 249 TDELHTYFNVNYTMHYLLNNGATDKLVMGVPFYGRAWSIEDRSKLGDPAGKMSPPGF 308 | QY | 246 PDETDELHTYFNVNYTMHYLLNNGATDKLVMGVPFYGRAWSIEDRSKLGDPAGKMS 305 |
| DB | 242 -PESDKILNDVYAINWIKNGTPKNKVLGMGYGSRFTLANAANGLGNAITGPSAGP 300 | DB | 880 PDDFDK--FNANFTIHWIEKGADPKLVGMGPMYQSPSLADNKEHGLNAKTYGGE 936 |
| QY | 309 ISGEGVLSYELCOLFQKEWHIOYDEYNAPYGYNDKIWGYDGLASISCKLAPLKL 368 | QY | 306 PGFISGEGVLSYELCOLFQKEWHIOYDEYN-APYGYNDKIWGYDGLASISCKLAP 364 |
| DB | 301 LTKPEGLMGVNEIC--SDRGWNEVFVEKAPYAYKQWGYDSVKSIGIKVDYLIRE 357 | DB | 937 AGESTRARGFLSYEICANIRKNTKWTARDKRGMGYAYKQWGYDGLASISCKLAP 996 |
| QY | 369 GVSQVWWSLENDDFKHC-CPKNPLLNKVNMMINGDEKNSFCILGPSTPTPTPT 427 | QY | 365 LKELGVSWWWSLENDDFKHC-CPKNPLLNKVNMMINGDEKNSFCIL--GPSTPTPT 421 |
| DB | 358 GLGGMTWLSLETDFRNCGGKYPPLTTIASKLNGD-----VARPTDPDK 403 | DB | 997 VKAMGLGGAMIWALDDFRNLDCCEEYPLRTINRLNRYNPGPGRCVLEKEPQREPR 1056 |
| QY | 428 TPTT 431 | QY | 422 PTTTPT 481 |
| DB | 404 QPTT 407 | DB | 1057 PTPRPTS-----TTPETTRRPTSTTTRRTTMTT---TTTRPTPTPTPTPTPT 1104 |
| RESULT 4 | | | |
| ID | 017412 PRELIMINARY; PRT; 1635 AA. | QY | 482 PSPTTTEHTSETPKY-----TTVVDGHLIKCYKEGDIPHPHTNHIKLVCEFNWGW 532 |
| AC | 017412; 01-JAN-1998 (TrEMBLrel. 05, Created) | DB | 1105 PTTTPT 1150 |
| DT | 01-JAN-1998 (TrEMBLrel. 05, Last sequence update) | QY | 533 WYHIMPCPPGTIW-----CQEKLT 551 |
| DT | 01-OCT-2003 (TrEMBLrel. 25, Last annotation update) | DB | 1151 -----LCPGGLYWSVDHCDWCPQSTNCRNKQT 1176 |
| DE | Probable chitinase 2 (EC 3.2.1.14). | RESULT 5 | |
| GN | CHT2. | Q9W223 | |
| OS | Aedes aegypti (Yellow fever mosquito). | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | |
| OC | Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes. | | |
| NCBI_TaxID | 7159; | | |
| RX | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RX | MEDLINE=98324849; PubMed=9662472; | | |
| RA | de la Vega H., Specht C.A., Liu Y., Robbins P.W.; | | |
| RT | "Chitinases are a multi-gene family in Aedes, Anopheles and | | |
| RT | Drosophila."; | | |
| RL | Insect Mol. Biol. 7:233-239(1998). | | |
| CC | -!- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N- | | |
| CC | ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN. | | |
| CC | -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL | | |
| CC | HYDROLASES). | | |
| DR | EMBL; AF026492; AAB81850.1; -. | | |
| DR | PIR; T14075; T14075. | | |

```

ID Q9W2Z3 PRELIMINARY; PRT; 4498 AA.
AC Q9W2Z3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG2989 protein.
DE CG2989.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.E., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.M., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003449; AAF46544.1; -.
DR Flybase; FBgn0030171; CG2989.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CSM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco 18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR GlycoSIDase; Hydrolase
SQ SEQUENCE 4498 AA; 493096 MW; ABE5CE0AAB25489 CRC64;
```

```
DR InterPro: IPR001223; Glyco_hydro_18.
DR InterPro: IPR001579; Glyco_hydro_18AS.
DR Pfam: PF01607; CBM_14; 1.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR ProDom: PD000471; Glyco_hydro_18; 1.
DR SMART: SM00494; ChtBD2; 1.
DR SMART: SM00636; Glyco_18; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
DR Glycosidase: Hydrolase.
SQ SEQUENCE 544 AA; 61069 MW; 89872DFC1DA23753 CRC64;

Query Match 27.3%; Score 847; DB 5; Length 544;
Best Local Similarity 36.5%; Pred. No. 5.2e-45;
Matches 207; Conservative 78; Mismatches 218; Indels 64; Gaps 19;

QY 1 MNTIYAILSIMA-CIGLMNASIKRDNHYSKPMRIVCVGTWSVYHK-VDPYTTIEDIDP 58
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 1 MRAIFATLAVLASCAALVQCADSR-----ARIVCFNSWAVRPGVGRYGIEDIPV 51
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 FKCTHLMYGFADKIDYKIQVDPYQDDHNSWEKRGYERFNNLRLKNPELTMTISLGG 118
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 52 DLCTHLIYFVIGTEKSESVLIIDPELD-----VDKSGFRNFTSLRSKHPDKFWMVAVGG 106
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 WYEGSEKYSDMAANPYRQOFIOSVLDLQYKFDGLDLDWEYVPSGR-LGNPKIDKQNYL 177
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 107 WAEGGSKYSNVAQKSTRFSIKSVDFLKDYDFGLDLDWEYFGAADRGGFSFKDEFL 166
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 ALVRELKDAF--EPHGYLLTAASPGKDKIDRAYDIKELNKLFDWMNVNTYDHGWNF 235
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 167 YFQELKRAFTRADRGWELTAAPLANFRLMEGYHVPCLQELDAIHVMSYDLRGWAGF 226
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 YGNAPLYKRPDDELTHTYFNVNYHLYNNGATDKLVGMVPPYGRAMSID----- 290
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 227 ADVHSPLYKRPD-QWAYEKLANVDGNLWEKGCPTNKLWGIPTPYGRSFTLSAGNNY 285
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 291 -----RSKLKLDPAKGMSPPFISGEEGLSVIELCQLPQKE--EWHIQVDEYNAPY 342
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 286 GLGTYINKEAGGDPAFYNTATGF-----WAYEICTVEDADGSGTWKKWDEFKQCPY 338
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 GYNDKIVGWYDDLASTCKLAFKELGVSQWVWSLENDDFKHGCPKPNLKNVHMIN 402
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 339 AYGKTQVGVEDPRSEIVEIKNWKIEKGYLGAMTAMDMDFKGLCGEENFLIKLH---- 394
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 403 GDEKNSPECILGFSST---TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 459
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 395 ---KMSVTVVPARTGHAPTPEW--ARPPSTPSDSEGDPI---PTTTTIVKPTTT 445
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 460 TTPSPPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 518
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 446 RTTARPT--TTTIVKPHGTTEDFDINVRPEVEEPTTENEVDNADV-CNSEDYVDPKKE 502
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 519 IHKLVCFEYVNGGWVHIMCPGTIW 545
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 503 CSKYWRC--VNGEGVQ--FSCQPGTIF 525

RESULT 7
Q8MP05 ID Q8MP05 PRELIMINARY; PRT; 2838 AA.
AC Q8MP05;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CHIT5.
OS Tenebrio molitor (Yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE FROM N.A.
RA Royer V., Fraichard S., Bouhin H.;
```

```

Db      2499 VLDGSTWTLKPHDSW 2513

RESULT 8
Q8WR52
ID      Q8WR52      PRELIMINARY;      PRT;      566 AA.
AC
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Chitinase precursor (EC 3.2.1.14).
GN      CHIB4.
OS      Bombyx mori (Silk moth).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC      Bombycidae; Bombyx.
OX      NCBI_TaxID=7091;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Kinsu x Showa; TISSUE=Integument;
RA      Abdel-Banat B.M.A., Koga D.;
RT      "Alternative mRNA splicing Generates heterogeneity within Bombyx mori
RL      gene for chitinase.";
RL      Submitted (DRC-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF455139; AAL51080.1;
DR      GO; GO:0005179; C:extracellular; IEA.
DR      GO; GO:0008061; C:chitin binding; IEA.
DR      GO; GO:0008061; F:endochitinase activity; IEA.
DR      GO; GO:0008061; F:hydrolyase activity; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      GO; GO:0006030; P:chitin metabolism; IEA.
DR      InterPro; IPR002557; Chitin bind PerA.
DR      InterPro; IPR001223; Glyco_hydro_18.
DR      InterPro; IPR001579; Glyco_hydro_18AS.
DR      Pfam; PF01607; CBM_14; 1.
DR      Pfam; PF00704; Glyco_hydro_18; 1.
DR      ProDom; PD000471; Glyco_hydro_18; 1.
DR      SMART; SM00494; ChitBD2; 1.
DR      SMART; SM00636; Glyco_18; 1.
DR      PROSITE; PS01095; CHITINASE_18; 1.
KW      Signal; Hydrolase; Glycosidase.
FT      SIGNAL 21 POTENTIAL.
FT      CHAIN 22 566 MW; 43FC717B3F6917D0 CRC64;
SQ      SEQUENCE 566 AA; 63465 MW; 43FC717B3F6917D0 CRC64;

Query Match      27.0%; Score 839; DB 5; Length 566;
Best Local Similarity 35.9%; Pred. No. 1.7e-44;
Matches 207; Conservative 78; Mismatches 209; Indels 82; Gaps 20;

Qy      1 MKTYAIIISMA-CIGLMNASIKRHDNDYKQNPRIYCVYGTWSVYHK-VDPYTIIDIP 58
Db      1 MRAIFATLAVLASCAALVQCADSR-----ARIVCYFSNNAVYRPGVGRYGIEDIPV 51

Qy      59 FKCTHLYGFAKIDYKTYQVFPYQDDNHSWEKGEYFNNLRKNBELTNTISLGG 118
Db      52 DLCTHLYSYFGVTEKSESLVLIIDPELD-----VDKSGFRNFTSLRKHDPVFMVAVGG 106

Qy      119 WYEGSEKYSMAAAPTQYQIQSVLDFLOEYKFDGLDWEYVPGSR-LGNPKIDKQNYL 177
Db      107 WAEGGSKYSNVAQKSTEMFIRSVDFLKIDFDGLDWEYVPGADRGGSSDKDKFL 166

Qy      178 ALVRELKDAF--EPHGYYLLTAASFGKIDRAYDIKELNKLFDWMNMTYDHYGGWENF 235
Db      167 YFVQELKRAFIKRAAGRWELTAAPLANFRLMEGHVVELQELDAIHVMYDRLGNWAGF 226

Qy      236 YGHNAPLYKRPDETHLYTFNNVNTYHYLNNGATDKLVMPVFFYGRAWSIED----- 290
Db      227 ADVHSPLYKRPD--QWAYEKLNVNDGLNWEKGCPTNKLWVGIPFYGRFTLSAGNNY 285

Qy      291 -----RSKULGDPKAPKMPGPGFISGEEGVLSYIELCOLFOKE--EWHIQYDEYNAPY 342
Db      286 GLGTYYNKEAGGDPAPYTNATGF-----WAYEICTEVDADGSGWTCKWDFGKCPY 338

```

```

Qy      343 GYNDKIWVGYYDLASISCKLAFKELGVSGVWWSLENDKFGHCGPKNPLINKVHNMN 402
Db      339 AYKGTQWGYEDPRSRVEIKMNWIKEGYLGATWALDMDDFKGLGGEENFLIKLLHKMS 398

Qy      403 GDEKNSFECILQPSSTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 458
Db      399 -----TIVTPPARTGHTTTTPEWARP-PTSPSDPSEGDPIPTTTTIVTKPTT 444

Qy      459 PTTTSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 510
Db      445 TTTTARPTT-----TTTKVPHGTTTEDFDINVRPEVEELPT-ENEVDNADY-CNSE 493

Qy      511 GD-IPHTNINIKYLVCFVNGGWWHIMPCCPGTIW 545
Db      494 DDYVDPKCECKYKWC--VNGEGVQ--FSCQPGIIF 525

RESULT 9
O15993
ID      O15993      PRELIMINARY;      PRT;      467 AA.
AC      O15993;
DT      01-JAN-1998 (TREMBlrel. 05, Created)
DT      01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Pichi-3.
OS      Penaeus japonicus (Kuruma prawn).
OC      Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC      Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC      Penaeidae; Marsupenaeus.
OX      NCBI_TaxID=27405;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Watanabe T., Kono M., Aida K., Nagasawa H.;
RT      "Purification and molecular cloning of a chitinase expressed in the
RT      hepatopancreas of the penaeid prawn Penaeus japonicus.";
RL      Biochim. Biophys. Acta 0:0-0(1997).
DR      EMBL; AB008027; BAA22854.1; -.
DR      HSSP; P07254; ICTN.
DR      GO; GO:0005576; C:extracellular; IEA.
DR      GO; GO:0008061; F:chitin binding; IEA.
DR      GO; GO:0016787; F:hydrolase activity; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      GO; GO:0006030; P:chitin metabolism; IEA.
DR      InterPro; IPR002557; Chitin bind PerA.
DR      InterPro; IPR001223; Glyco_hydro_18.
DR      InterPro; IPR001579; Glyco_hydro_18AS.
DR      Pfam; PF01607; CBM_14; 1.
DR      Pfam; PF00704; Glyco_hydro_18; 1.
DR      ProDom; PD000471; Glyco_hydro_18; 1.
DR      SMART; SM00494; ChitBD2; 1.
DR      SMART; SM00636; Glyco_18; 1.
DR      PROSITE; PS01095; CHITINASE_18; 1.
SQ      SEQUENCE 467 AA; 51765 MW; 499F7095774CA445 CRC64;

Query Match      27.0%; Score 837.5; DB 5; Length 467;
Best Local Similarity 34.8%; Pred. No. 1.7e-44;
Matches 183; Conservative 85; Mismatches 173; Indels 85; Gaps 14;

Qy      35 IVCYGTWSVYHK-VDPYTIIDIPKCTHLYMGFAKIDYKTYQVFPYQDDNHS 91
Db      1 MVCYFGSAVYRQGLGKEDVEDIDPKICTHIVFGFAGL-AADSSIRVLDPNLCNDYTG- 58

Qy      92 WEKRGYERFNNLRKNBELTNTISLGGWYEGSEKYSMAAAPTQYQIQSVLDFLOEYK 151
Db      59 --KCAVDYRFTALKQQAANLKALLAYGVNWSGPKYSKVAADPALNRRFITSSIELKKHG 116

Qy      152 FDGLDLDEYPCSGNCPKIDKQNYLALVRELKDAFEPHGYYLLTAASFGKIDRAYDI 211
Db      117 FDGLDMDWEYFQRCGSPD-DYDNFALIMAEKQALQPEGLMLTAAVAGKATIDPATNV 175

Qy      212 KLANLFDWMNMTYDHYGGWENFYGHNAPLYKRP-DETDELHTYFNNVNTYHYLNNGA 270
Db      176 PEISKSLDLINWSDYDLHGAMDYTHHQSGLYAHPLDSEGNL--YLVNDFALSYWIEKA 233

```



```

QY 271 TRDKLVGVPPVYGRAWSIEDRSKLGDPAGKMSPPGFISBEGVLSVIELCOLFQKEW 330
Db 234 RFGQIALGIPLYGRWCMTLASQOETGYAPAHQFGAAGDWTKSPGMLGYNEICYMOTTDW 293
QY 331 HIQVDEYNAPYGY--NDKIWGVDDDLASISCKLAFKLGLGVSGVMWSLENDDFKGHC 387
Db 294 TVDDPAMHEPYYAYFPNNIWCSDHAASVVTKAEYAKSKGLAGMWSVETDDFRGLC 353
QY 388 GPKNLLNKNVNMINGDEKNSPECILGPSTTTPTTTPTTTPTTTPTTTPTTTPTTPTSP 447
Db 354 H-----NRKYHLI---KTMVEVFGSGSITEP----- 376
QY 448 TPTTTPTSPPTTTPTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTT 507
Db 377 -----PPLPTTTRDPSEP---TTTRAPP-----EG-----IHC 403
QY 508 YKEGDIPHPTNIHKLVCEP-VNGGWWVHIMCPPEGTIWCQEKLTG 552
Db 404 TTLGLNPDLDTHYLCSLNTSGGFEKEEYCPREGTLFNPQSFYC 449

RESULT 10
Q9GR93 PRELIMINARY; PRT; 543 AA.
AC Q9GR93;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CHIB.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kinshu x Showa; TISSUE=Integument;
RA Abdel-Banat B.M., Koga D.;
RT "Molecular cloning of Bombyx mori chitinase cDNA: a unique insert of 9
RT base pairs reduced the apparent molecular mass of the encoded
RT protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB052914; BAB20017.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0008843; F:endochitinase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 543 CHITINASE.
SQ SEQUENCE 543 AA; 60971 MW; C36C71E3B84DF8F8 CRC64;

Query Match 27.0%; Score 837.5; DB 5; Length 543;
Best Local Similarity 35.9%; Pred. No. 2.1e-44;
Matches 207; Conservative 78; Mismatches 208; Indels 83; Gaps 20;

Qy 1 MKTYAIIISWA-CIGLNNASIKRDNHYSKNPMEIVCYVGTWSYTHK-VDPTYTTEDIDP 58
Db 1 NRAIFATLAVASCAALVQ-----SDSRARIVCYFSNWAYRSGVGRYGEDIPV 50

```


OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003701; AAF54987.1;
DR EMBL; AY061553; AAL29101.1;
DR FlyBase; FBGN0038180; CG9307.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind perA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChCBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR GlycoSIDe; Hydrolase.
SQ SEQUENCE 595 AA; 67050 MW; 8DC8469D3F732593 CRC64;

Query Match 27.0%; Score 837.5; DB 5; Length 595;
Best Local Similarity 35.2%; Pred. No. 2.3e-44;
Matches 201; Conservative 84; Mismatches 199; Indels 87; Gaps 22;
QY 29 SKPMRIVCYGVTSVYHK-VDPYTTIEDIDFFKTHLMYGFADKIDYKTYQVDFPDQDD 87
DB 23 SDQASRIVCYFSNVAVTGTIGRYGLEDVPADLCTHIYFISGVNDKSWDLVIDPDL- 81
QY 88 NHNSWEKRGVERFNNLKLKPELTWISLGGWTEGSEKYSMDAANTYRQQTQSVDLDEL 147
DB 82 ----VDQGSFKEFTQLKSNPNVLEIAVGMWAGGSKYSQMVAVRRDRRSFIRSVRFM 137
QY 148 QEYKFDGLDLDWEYVGS--FLGNPKDKQNYALVRELKDAF--EPHYLLTAAVSPGKD 203
DB 138 KQYNFDDFDLDWEYVGTATDGGYNG-DKDKELFVLELRAPDREGREWITWAVPAKF 196
QY 204 KIDRAYDIKELNKLFDMMNVTYDYHGGWENFYGHNAPLYKRPDETDLHLYNVNVTMH 263
DB 197 RLNEGYHVPCLCEALDAIHAMTYDLRGNWAGFADVHSPLYKPKHD-QYAYEKLNVNDGLA 255
QY 264 YYLANGATROKLVGVFPFYGRAMSIEDRSK-LKLG----DPAKMSPPGFISGEGVLSYI 319
DB 256 LWEEMGCPANKLVGVFPFYGRFTFTLSNKNYNNMTYINKEAGGGAPGPYTNASGFLAYY 315
QY 320 ELC-QLFKQEE-WHIQYDEVYNAPYGYNDKLVGYDDLASICKLAFKELGVGVWVWS 377
DB 316 EICTEVNDKSGMTWDDAGWVPYTYKOTQWGVYENASIQKDFIKORGAGATWTA 375
QY 378 LENDDPKGHGCPKNFLANKVH-NMINGDEKNSPECILGPSTTTTPTTTP-----TTPTTT 432
DB 376 IDMDDFHGMCGRKNGLTQILDNMKN-----YRVEPTRQTTTRPEWAKXPAT 423
QY 433 PTTP-----SPTPTT-----TPSPTTPTTTPS--PTTPTTPTTPTTPTTPTT-- 473
DB 424 PPNDEGANVAPTSTTKRKPCKPSTSSPLSPSAPGPVTVGSSTPKPTTKKPKPKK 483
QY 474 TPAPTTPSTPTTTEHTSETPKYTY-----VDGLIKCYKEGDIPHP 516
DB 484 TTTTITTTTAPD---EKSTEEPEEVYVDPVEPTDPEQPMGPQDPDNEIDCTNRDFVPH 540
QY 517 TNIHKYLVCFFVNGWVWHNP-----CPPGT 543
DB 541 -NCRKYFR-----VHGKPEVFECKEGT 562

RESULT 14

Q9W2M7 PRELIMINARY; PRT; 460 AA.

AC Q9W2M7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG9357 protein.
GN CG9357.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weistock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
DR EMBL; AE003452; AAF46663.1; --
DR HSSP; P07254; ICTN.
DR FlyBase; FBgn0034580; CG9357.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001179; FKBP Pfase.
DR InterPro; IPR001223; Glyco hydro 18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 460 AA; 51055 MW; 8EP4F126F1D8D9B CRC64;

Query Match 26.8%; Score 832; DB 5; Length 460;
Best Local Similarity 33.7%; Pred. No. 3.7e-44;
Matches 183; Conservative 87; Mismatches 175; Indels 98; Gaps 11;

QY 17 KNASIKRHDNDYKSNPMRIVCVGWSVYHK-VDPYTTIEDIPFKCTHLMYGFPAKIDYK 75
DB 1 MAASSAAGNS-SKN---VVCQGWSTVRPGLGKFGMEDIDPFCTHLYAFGLIEE-T 55

QY 76 YTIQVFDYQDDNHNHNSWEKRGYERFNNLRKLNPELTMIISLGWYEGSEKYSDMAANPTY 135
DB 56 GOLRVDAVLDLEENS-GEGNKSFNALKKNPVLKTLVAVGWNEGSKRFSLVARDPSK 114

QY 136 RQQTQSVDLFQEQYKFGDLDDWEYPGSRGNPKIKQNYALVRELKDAFEPHGYLLT 195
DB 115 REKFVDDVVRFLQRFQFGDLDDWEYPGSRHSLDNEGRSNITFLKELKEGEPFGFILS 174

QY 196 AAVSPGKDKIDRAYDIKELNKLFDWNNYTYDYGWENFYGHNAPLY---KRPDETEL 252
DB 175 AAVGAQFAEISYDIPAMVPYLDLINVWAYDLHGFWQVQGINAPLYAAEKDASDSGR 234

QY 253 HYFVNNYTHVYLNNGATRDKLVGVPYFGRAWGIEDRSKLKLGDPKAGMSPPGFISGE 312

DB 235 QQQLNDVAVKYLKAGAPAEKILGVFFYGRSFTLTATAGNPGAPGHICKIAGNYSRE 294
QY 313 EGVLSYTELCQFOKEWHIYQDEYVNPAGYNDKIWVGYYDDLASISCKLAFKELGVSG 372
DB 295 PGVLGYNELCEMREERTQKWEATQVPAYRORQWGVEDPRSLAKAQYVMDNLGG 354
QY 373 VVWVSLNDDFKGCGPK-NPLNKNVHNMINGDEKNSFECILGPSTTTPTTTPTTTPTTT 431
DB 355 IMVWLSLEDDFRGTCGQPPYELLHEINRVLFGG-----NTPSGLTTSNR 399
QY 432 TPTTPT 491
DB 388 -----NTPSGLTTSNR 399
QY 492 ETPK--YTTYVDGHLIKCYKEGDIHPHTNIHKLVCFVNGVWVHMPCPGPIWCOEK 549
DB 400 ESPSGFGCPADA-----PAGYIRDNDCKSFYIC-----SGGKTHNFDCEGLNFDLDT 449
QY 550 LTC 552
DB 450 KSC 452

RESULT 15
Q8ITU3 PRELIMINARY; PRT; 467 AA.
AC Q8ITU3; (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase.
OS *Penaeus vannamei* (Penaeid shrimp) (European white shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Litopenaeus.
OX NCBI_TaxID=6689;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang R.C.;
RT "Cloning and characterization of a cDNA encoding a chitinase from
RT hepatopancreas of the *Penaeus vannamei* (Crustacea, Decapoda).";
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF313689; AAN74647.1; --
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
SQ SEQUENCE 467 AA; 51958 MW; 734A830C6F47F4CD CRC64;

Query Match 26.8%; Score 831.5; DB 5; Length 467;
Best Local Similarity 34.2%; Pred. No. 4.1e-44;
Matches 180; Conservative 87; Mismatches 174; Indels 85; Gaps 14;

QY 35 IVCYVGTWVYHK-VDPYTTIEDIPFKCTHLMYGFPAKIDYKTYQVFPYOD--DNHNS 91
DB 1 MVCFYGSAAVYRGLGKFDVEDIDPKCTHIFGFLAGL-AHDSIRVLDPWNLCDNYG- 58

QY 92 WEKRGYERFNNLRKLNPELTMIISLGWYEGSEKYSDMAANPTYRQQTQSVDLFQEQYK 151
DB 59 --KCAVDRFTALQKQANLKAIIAVGWNEGSPKYSKVAADPYLRNRFITSSLELKKHG 116

QY 152 FDGLDDWEYPGSRGNPKIKQNYALVRELKDAFEPHGYLLTAAVSPGKDKIDRAYDI 211

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:59:34 ; Search time 4.09116 Seconds
(without alignments)
1265.926 Million cell updates/sec

Title: US-09-662-293-13
Perfect score: 105
Sequence: 1 DPAKGMPPGFIVGEEGVLS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------------|-------------------|
| 1 | 105 | 100.0 | 20 | 14 US-10-218-743-4 | Sequence 4, Appli |
| 2 | 105 | 100.0 | 20 | 14 US-10-218-743-13 | Sequence 13, Appl |
| 3 | 101 | 96.2 | 490 | 14 US-10-218-743-41 | Sequence 41, Appl |
| 4 | 101 | 96.2 | 509 | 14 US-10-218-743-35 | Sequence 35, Appl |
| 5 | 101 | 96.2 | 509 | 14 US-10-218-743-38 | Sequence 38, Appl |
| 6 | 99 | 94.3 | 536 | 14 US-10-218-743-21 | Sequence 21, Appl |
| 7 | 99 | 94.3 | 555 | 14 US-10-218-743-15 | Sequence 15, Appl |
| 8 | 99 | 94.3 | 555 | 14 US-10-218-743-18 | Sequence 18, Appl |
| 9 | 50 | 47.6 | 143 | 12 US-10-424-599-238601 | Sequence 238601, |
| 10 | 49 | 46.7 | 69 | 12 US-10-424-599-181344 | Sequence 181344, |
| 11 | 48 | 45.7 | 57 | 9 US-09-925-302-483 | Sequence 483, App |
| 12 | 48 | 45.7 | 350 | 13 US-10-067-615-6 | Sequence 6, Appli |
| 13 | 48 | 45.7 | 350 | 13 US-10-076-754-6 | Sequence 6, Appli |
| 14 | 48 | 45.7 | 350 | 13 US-10-076-773-6 | Sequence 6, Appli |
| 15 | 48 | 45.7 | 411 | 8 US-08-916-625B-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1

```
US-10-218-743-4
; Sequence 4, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-4
```

Query Match 100.0%; Score 105; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPAKGMPPGFIVGEEGVLS 20

Db 1 DPAKGMPPGFIVGEEGVLS 20

```
16 48 45.7 411 9 US-09-828-739-2 Sequence 2, Appli
17 48 45.7 411 9 US-09-757-421-4 Sequence 4, Appli
18 48 45.7 411 9 US-09-874-138-2 Sequence 2, Appli
19 48 45.7 411 9 US-09-887-879-11 Sequence 11, Appli
20 48 45.7 411 9 US-09-811-088-6 Sequence 6, Appli
21 48 45.7 411 9 US-09-992-964-11 Sequence 11, Appli
22 48 45.7 411 12 US-10-211-462-67 Sequence 67, Appli
23 48 45.7 411 13 US-10-039-785-3 Sequence 3, Appli
24 48 45.7 411 13 US-10-005-842-2 Sequence 2, Appli
25 48 45.7 411 13 US-10-067-615-2 Sequence 2, Appli
26 48 45.7 411 13 US-10-052-798-1 Sequence 1, Appli
27 48 45.7 411 13 US-10-076-754-2 Sequence 2, Appli
28 48 45.7 411 13 US-10-076-773-2 Sequence 2, Appli
29 48 45.7 411 14 US-10-207-655-194 Sequence 194, App
30 48 45.7 411 14 US-10-314-410-6 Sequence 6, Appli
31 48 45.7 411 14 US-10-242-383-11 Sequence 11, Appli
32 48 45.7 411 14 US-10-288-917-1 Sequence 1, Appli
33 48 45.7 411 14 US-10-322-673-3 Sequence 3, Appli
34 48 45.7 411 14 US-10-139-785-3 Sequence 3, Appli
35 48 45.7 411 15 US-10-292-486-3 Sequence 3, Appli
36 48 45.7 411 15 US-10-333-712-4 Sequence 4, Appli
37 48 45.7 411 15 US-10-423-448-1 Sequence 1, Appli
38 48 45.7 440 9 US-09-757-421-2 Sequence 2, Appli
39 48 45.7 440 9 US-09-811-088-4 Sequence 4, Appli
40 48 45.7 440 14 US-10-314-410-4 Sequence 4, Appli
41 48 45.7 440 15 US-10-333-712-3 Sequence 3, Appli
42 46 43.8 34 12 US-10-424-599-169740 Sequence 169740,
43 46 43.8 135 12 US-10-424-599-257787 Sequence 267787,
44 45 42.9 123 12 US-10-425-114-57295 Sequence 57295, A
45 45 42.9 303 9 US-09-738-626-6409 Sequence 6409, Ap
```

```
RESULT 2
US-10-218-743-13
; Sequence 13, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-13

Query Match          100.0%; Score 105; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEGVLS 20
| | | | | | | | | | | | | | | | | | | |
DB 1 DPAKGMSPPGFIVGEGVLS 20

RESULT 3
US-10-218-743-41
; Sequence 41, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-41
```

```
Query Match          96.2%; Score 101; DB 14; Length 490;
Best Local Similarity 95.0%; Pred. No. 9.4e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEGVLS 20
| | | | | | | | | | | | | | | | | | | |
DB 279 DPAKGMSPPGFITGEGVLS 298

RESULT 4
US-10-218-743-35
; Sequence 35, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-35

Query Match          96.2%; Score 101; DB 14; Length 509;
Best Local Similarity 95.0%; Pred. No. 9.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEGVLS 20
| | | | | | | | | | | | | | | | | | | |
DB 298 DPAKGMSPPGFITGEGVLS 317

RESULT 5
US-10-218-743-38
; Sequence 38, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-38

Query Match          96.2%; Score 101; DB 14; Length 509;
Best Local Similarity 95.0%; Pred. No. 9.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEEGVL 20
Db 298 DPAKGMSPPGFIVGEEGVL 317

RESULT 6
US-10-218-743-21
; Sequence 21, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-21

Query Match          94.3%; Score 92; DB 14; Length 536;
Best Local Similarity 95.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEEGVL 20
Db 279 DPAKGMSPPGFIVGEEGVL 298

RESULT 7
US-10-218-743-15
; Sequence 15, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
```

```
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-15

Query Match          94.3%; Score 99; DB 14; Length 555;
Best Local Similarity 95.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEEGVL 20
Db 298 DPAKGMSPPGFIVGEEGVL 317

RESULT 8
US-10-218-743-18
; Sequence 18, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-18

Query Match          94.3%; Score 99; DB 14; Length 555;
Best Local Similarity 95.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEEGVL 20
Db 298 DPAKGMSPPGFIVGEEGVL 317

RESULT 9
US-10-424-599-238601
; Sequence 238601, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David X
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 238601
; LENGTH: 143
; TYPES: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57482C.1.pep
US-10-424-599-238601

Query Match 47.6%; Score 50; DB 12; Length 143;
Best Local Similarity 61.1%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 PAKGMSPPGF--IVGEEG 17
Db 121 PNKGDPPKFFNVGEEG 138

RESULT 10
US-10-424-599-181344
; Sequence 181344, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181344
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_134769C.1.pep
US-10-424-599-181344

Query Match 46.7%; Score 49; DB 12; Length 69;
Best Local Similarity 72.7%; Pred. No. 6.8;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PAKGMSPPGFI 12
Db 31 PKEGVSPPGFI 41

RESULT 11
US-09-925-302-483
; Sequence 483, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 483
; LENGTH: 57
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-925-302-483
Query Match 45.7%; Score 48; DB 9; Length 57;
Best Local Similarity 57.1%; Pred. No. 7.8;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 MSPPGFIVGEEV 19
Db 9 LSPGFWLGVGIL 22

RESULT 12
US-10-067-615-6
; Sequence 6, Application US/10067615
; Publication No. US20020115154A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL TRAIL RECEPTORS, NUCLEIC ACIDS ENCODING SAME, AND
; FILE REFERENCE: 480140.432D1
; CURRENT APPLICATION NUMBER: US/10/067,615
; CURRENT FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-615-6

Query Match 45.7%; Score 48; DB 13; Length 350;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 PAKGMSPPGFIVGEEG 17
Db 76 PSEGLCPFGHISEDG 91

RESULT 13
US-10-076-754-6
; Sequence 6, Application US/10076754
; Publication No. US20020161195A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL TRAIL RECEPTORS, NUCLEIC ACIDS ENCODING SAME, AND
; FILE REFERENCE: 480140.432D3
; CURRENT APPLICATION NUMBER: US/10/076,754
; CURRENT FILING DATE: 2002-02-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-754-6

Query Match 45.7%; Score 48; DB 13; Length 350;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 PAKGMSPPGFIVGEEG 17
Db 76 PSEGLCPFGHISEDG 91

RESULT 14
US-10-076-773-6
; Sequence 6, Application US/10076773
; Publication No. US20020161196A1
; GENERAL INFORMATION:


```

Matches      7; Conservative      5; Mismatches      4; Indels      0; Gaps      0;

Qy      2      PAKGMSPPGFIVCEEG      17
      ||::|||::|
Db      76      PSEGLCPGHHISEDG      91

Search completed: March 22, 2004, 07:45:49
Job time : 4.09116 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:42:54 ; Search time 1.61201 Seconds
(without alignments)
640.518 Million cell updates/sec

Title: US-09-662-293-13

Perfect score: 105

Sequence: 1 DPAKGMSPPGFIVGEEGVLS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgm2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgm2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgm2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgm2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgm2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 105 | 100.0 | 20 | 4 | US-09-292-225-4 |
| 2 | 105 | 100.0 | 20 | 4 | US-09-292-225-13 |
| 3 | 101 | 96.2 | 490 | 4 | US-09-292-225-41 |
| 4 | 101 | 96.2 | 509 | 4 | US-09-292-225-35 |
| 5 | 101 | 96.2 | 509 | 4 | US-09-292-225-38 |
| 6 | 99 | 94.3 | 536 | 4 | US-09-292-225-21 |
| 7 | 99 | 94.3 | 555 | 4 | US-09-292-225-15 |
| 8 | 99 | 94.3 | 555 | 4 | US-09-292-225-18 |
| 9 | 48 | 45.7 | 350 | 4 | US-09-134-618-6 |
| 10 | 48 | 45.7 | 411 | 3 | US-09-329-613A-2 |
| 11 | 48 | 45.7 | 411 | 4 | US-09-079-029-1 |
| 12 | 48 | 45.7 | 411 | 4 | US-09-134-618-2 |
| 13 | 48 | 45.7 | 411 | 4 | US-10-039-785-3 |
| 14 | 48 | 45.7 | 412 | 4 | US-09-333-593A-2 |
| 15 | 48 | 45.7 | 424 | 4 | US-09-333-593A-8 |
| 16 | 48 | 45.7 | 440 | 3 | US-08-883-036A-2 |
| 17 | 48 | 45.7 | 440 | 4 | US-09-536-201-2 |
| 18 | 48 | 45.7 | 440 | 4 | US-09-578-392-2 |
| 19 | 47 | 44.8 | 420 | 4 | US-09-252-991A-24263 |
| 20 | 45 | 42.9 | 503 | 4 | US-09-252-991A-30682 |
| 21 | 44 | 41.9 | 219 | 4 | US-09-252-991A-20817 |
| 22 | 44 | 41.9 | 3224 | 2 | US-08-705-660-34 |
| 23 | 44 | 41.9 | 3224 | 3 | US-08-989-045-34 |
| 24 | 42 | 40.0 | 411 | 4 | US-09-252-991A-18602 |
| 25 | 41.5 | 39.5 | 915 | 4 | US-09-543-681A-5334 |
| 26 | 41 | 39.0 | 367 | 4 | US-09-252-991A-16628 |
| 27 | 41 | 39.0 | 385 | 4 | US-09-134-000C-4952 |

RESULT 1

US-09-292-225-4
; Sequence 4, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-4

Query Match 100.0%; Score 105; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.6e-10; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPAKGMSPPGFIVGEEGVLS 20

Db 1 DPAKGMSPPGFIVGEEGVLS 20

RESULT 2

US-09-292-225-13
; Sequence 13, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225


```
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; SOFTWARE: Patentin Ver. 2.0
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
; US-09-292-225-21

Query Match          94.3%; Score 99; DB 4; Length 536;
Best Local Similarity 95.0%; Pred. No. 6.3e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DPAKGMSPPGFIVGEEGVLS 20
Db 279 DPAKGMSPPGFISGEEGVLS 298

RESULT 7
US-09-292-225-15
; Sequence 15, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; SOFTWARE: Patentin Ver. 2.0
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
; US-09-292-225-15

Query Match          94.3%; Score 99; DB 4; Length 555;
Best Local Similarity 95.0%; Pred. No. 6.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DPAKGMSPPGFIVGEEGVLS 20
Db 298 DPAKGMSPPGFISGEEGVLS 317

RESULT 8
US-09-292-225-18
; Sequence 18, Application US/09292225
```

```
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
; US-09-292-225-18

Query Match          94.3%; Score 99; DB 4; Length 555;
Best Local Similarity 95.0%; Pred. No. 6.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DPAKGMSPPGFIVGEEGVLS 20
Db 298 DPAKGMSPPGFISGEEGVLS 317

RESULT 9
US-09-134-618-6
; Sequence 6, Application US/09134618
; Patent No. 6417328
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL TRAIL RECEPTORS, NUCLEIC ACIDS ENCODING SAME, AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 480140.432
; CURRENT APPLICATION NUMBER: US/09/134,618
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-134-618-6

Query Match          45.7%; Score 48; DB 4; Length 350;
Best Local Similarity 43.8%; Pred. No. 6.7;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PAKGMSPPGFIVGEEG 17
Db 76 PSEGLCPGHHISDG 91

RESULT 10
US-09-329-633A-2
; Sequence 2, Application US/09329633A
; Patent No. 6252050
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chuntharapai, Anan
; APPLICANT: Kim, K. Jin
; TITLE OF INVENTION: METHOD FOR MAKING MONOCLONAL ANTIBODIES AND
; CROSS-REACTIVE ANTIBODIES OBTAINABLE BY THE METHOD
```

```
FILE REFERENCE: P1468R1 (REVISED)
CURRENT APPLICATION NUMBER: US/09/329,633A
CURRENT FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/089,253
PRIOR FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
LENGTH: 411
TYPE: PRT
ORGANISM: human
FEATURE:
NAME/KEY: xaa
LOCATION: 410
OTHER INFORMATION: xaa = leu or met
US-09-329-633A-2

Query Match          45.7%; Score 48; DB 3; Length 411;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 PAKGMSPPGFIVGEEG 17
Db      76 PSEGLCPGGHISEDG 91

RESULT 11
US-09-079-029-1
; Sequence 1, Application US/09079029
; Patent No. 6342369
; GENERAL INFORMATION:
; APPLICANT: Adams, Camilia W.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chundharapal, Anan
; APPLICANT: Kim, Kyung J.
; TITLE OF INVENTION: Apo-2 Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-079-029-1

Query Match          45.7%; Score 48; DB 4; Length 411;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 PAKGMSPPGFIVGEEG 17
Db      76 PSEGLCPGGHISEDG 91

FILE REFERENCE: P1468R1 (REVISED)
CURRENT APPLICATION NUMBER: US/09/329,633A
CURRENT FILING DATE: 1999-06-10
PRIOR APPLICATION NUMBER: US 60/089,253
PRIOR FILING DATE: 1998-06-12
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 2
LENGTH: 411
TYPE: PRT
ORGANISM: human
FEATURE:
NAME/KEY: xaa
LOCATION: 410
OTHER INFORMATION: xaa = leu or met
US-09-329-633A-2

Query Match          45.7%; Score 48; DB 3; Length 411;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 PAKGMSPPGFIVGEEG 17
Db      76 PSEGLCPGGHISEDG 91

RESULT 12
US-09-134-618-2
; Sequence 2, Application US/09134618
; Patent No. 6417328
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL TRAIL RECEPTORS, NUCLEIC ACIDS ENCODING SAME, AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 480140.432
; CURRENT APPLICATION NUMBER: US/09/134,618
; NUMBER OF SEQ ID NOS: 6
; NUMBER OF FILING DATE: 1998-08-14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-134-618-2

Query Match          45.7%; Score 48; DB 4; Length 411;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 PAKGMSPPGFIVGEEG 17
Db      76 PSEGLCPGGHISEDG 91

RESULT 13
US-10-039-785-3
; Sequence 3, Application US/10039785
; Patent No. 6538938
; GENERAL INFORMATION:
; APPLICANT: Salcedo et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind to TRAIL
; FILE REFERENCE: PF550
; CURRENT APPLICATION NUMBER: US/10/039,785
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 60/369,860
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/341,237
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/331,310
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/331,044
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: 60/327,364
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/323,807
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/309,176
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/294,981
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/293,473
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-039-785-3

Query Match          45.7%; Score 48; DB 4; Length 411;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      2 PAKGMSPPGFIVGEEG 17
Db      76 PSEGLCPGGHISEDG 91
```

Db 76 PSEGLCPPGHHISEDG 91

RESULT 14

US-09-333-593A-2

; Sequence 2, Application US/09333593A
; Patent No. 6313269
; GENERAL INFORMATION:
; APPLICANT: DEEN, KEITH C.
; APPLICANT: YOUNG, PETER R.
; APPLICANT: MARSHALL, LISA A.
; APPLICANT: ROSHAK, AMY K.
; APPLICANT: TAN, KONG B.
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RELATED RECEPTOR,
; TITLE OF INVENTION: TR6
; FILE REFERENCE: GH-50008-2
; CURRENT APPLICATION NUMBER: US/09/333,593A
; CURRENT FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 08/916,625
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/853,684
; PRIOR FILING DATE: 1997-05-09
; PRIOR APPLICATION NUMBER: 60/041,230
; PRIOR FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 412
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-333-593A-2

Query Match 45.7%; Score 48; DB 4; Length 412;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PAKGMSPPGFIVGEG 17

Db 76 PSEGLCPPGHHISEDG 91

RESULT 15

US-09-333-593A-8

; Sequence 8, Application US/09333593A
; Patent No. 6313269
; GENERAL INFORMATION:
; APPLICANT: DEEN, KEITH C.
; APPLICANT: YOUNG, PETER R.
; APPLICANT: MARSHALL, LISA A.
; APPLICANT: ROSHAK, AMY K.
; APPLICANT: TAN, KONG B.
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RELATED RECEPTOR,
; TITLE OF INVENTION: TR6
; FILE REFERENCE: GH-50008-2
; CURRENT APPLICATION NUMBER: US/09/333,593A
; CURRENT FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 08/916,625
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/853,684
; PRIOR FILING DATE: 1997-05-09
; PRIOR APPLICATION NUMBER: 60/041,230
; PRIOR FILING DATE: 1997-03-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 424
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-333-593A-8

Query Match 45.7%; Score 48; DB 4; Length 424;

Best Local Similarity 43.8%; Pred. No. 8.3;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PAKGMSPPGFIVGEG 17

Db 76 PSEGLCPPGHHISEDG 91

Search completed: March 22, 2004, 07:03:57
Job time : 1.61201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:40:28 ; Search time 36.0951 Seconds
(without alignments)
1479.047 Million cell updates/sec

Title: US-09-662-293-15
Perfect score: 3107
Sequence: 1 MKTIYAILSIMACIGLMNAS.....IMPCPPQTWCQKLTICIG 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|---------------------|
| 1 | 1008.5 | 32.5 | 525 | 2 T44445 | chitinase (EC 3.2.) |
| 2 | 869 | 28.0 | 214075 | 2 T14075 | chitinase (EC 3.2.) |
| 3 | 838.5 | 27.0 | 554 | 2 A56596 | chitinase (EC 3.2.) |
| 4 | 802 | 25.8 | 483 | 2 A53918 | chitinase (EC 3.2.) |
| 5 | 705.5 | 22.7 | 617 | 2 T15408 | hypothetical prote |
| 6 | 685.5 | 22.1 | 537 | 2 S57197 | oviduct-specific g |
| 7 | 685 | 22.0 | 539 | 2 I46470 | estrogen dependent |
| 8 | 668 | 21.5 | 654 | 2 I38605 | oviductal glycopro |
| 9 | 662 | 21.3 | 383 | 2 A49562 | cartilage glycopro |
| 10 | 650 | 20.9 | 383 | 2 S51327 | heparin-binding gl |
| 11 | 646.5 | 20.8 | 405 | 2 S61551 | breast-regressing |
| 12 | 608 | 19.6 | 399 | 2 S27879 | secretory protein |
| 13 | 604.5 | 19.5 | 504 | 2 A38221 | chitinase (EC 3.2.) |
| 14 | 592 | 19.1 | 599 | 2 D83764 | chitinase BH0916 [|
| 15 | 561 | 18.1 | 699 | 2 A38368 | chitinase (EC 3.2.) |
| 16 | 560 | 18.0 | 1215 | 2 T43916 | chitinase A [impor |
| 17 | 463.5 | 14.9 | 831 | 2 T00323 | chitinase (EC 3.2.) |
| 18 | 432.5 | 13.9 | 1484 | 2 T29275 | hypothetical prote |
| 19 | 428 | 13.8 | 869 | 2 T44440 | chitinase (EC 3.2.) |
| 20 | 427.5 | 13.8 | 849 | 2 D82510 | chitinase VCA0027 |
| 21 | 424.5 | 13.7 | 2025 | 2 T03884 | hypothetical prote |
| 22 | 414.5 | 13.3 | 563 | 2 S60651 | chitinase precursor |
| 23 | 413 | 13.3 | 756 | 2 AB1452 | chitinase B homolo |
| 24 | 412 | 13.3 | 423 | 2 JQ1975 | chitinase (EC 3.2.) |
| 25 | 411.5 | 13.2 | 424 | 2 S68121 | chitinase I precu |
| 26 | 411 | 13.2 | 423 | 2 S51369 | chitinase - fungu |
| 27 | 409.5 | 13.2 | 756 | 2 AB1088 | chitinase B homolo |
| 28 | 408.5 | 13.1 | 561 | 2 A25090 | chitinase (EC 3.2.) |
| 29 | 405 | 13.0 | 546 | 2 F84238 | chitinase [impor |

RESULT 1

T44445
Chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C/Species: Anopheles gambiae (African malaria mosquito)
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C/Accession: T44445
R/Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A/Reference number: Z22771
A/Accession: T44445
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-525 <SHE>
A/Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A/Experimental source: adult; gut
C/Genetics:
A/Gene: chi-1
C/Keywords: glycosidase; hydrolase

ALIGNMENTS

| | | | | | |
|-----------------------|-----|--|--------------------|------------|-------------|
| Query Match | | 32.5%; | Score 1008.5; | DB 2; | Length 525; |
| Best Local Similarity | | 38.4%; | Pred. No. 2.2e-55; | | |
| Matches 211; | | Conservative 93; | Mismatches 187; | Indels 59; | Gaps 13; |
| QY | 8 | LSIMACIGLMNASIKRD-HNDYSKNPMRIVCVYGTWSVYHKVD-PYTIEDIDPFKCTHLM | 65 | | |
| DB | 5 | VGVLVIVAFAAAFAEPEHKAASAEKKVCYGVGTWAVYRPGNGRYDIEHIDPSLCTHLM | 64 | | |
| QY | 66 | YGFAXIDEVYKTIQVDPYQDDNHNHWEKRGYERFNNLRKKNPELTWISLGGWEGSEK | 125 | | |
| DB | 65 | YGFPGINE-DATVRIIDPYLDLEEN-WGRGHIKRFVGLKNVGPGLKTLAAIGMWEGSRK | 122 | | |
| QY | 126 | YSDMAANPTVYRQQTQSVDLQEQYKFDGLDLDWEYVPGSRGLGNPKDKNYLALVRELKD | 185 | | |
| DB | 123 | FSAAMASGELRRKRFISDCVAFQCRHGFQIDLDWYPAQRDGNPLIDNDNHAQLVEEKR | 182 | | |
| QY | 186 | AFEPHGYLLTAASVSPGKDKIDRAYDIKELINKLFDWMNVMTVDYHGGWENFYGHNAFLYKR | 245 | | |
| DB | 183 | EDHYGLLLTAASVASVESAGVSVDIPRISKSFHLNVMTVDHMGANVSYCINAPLYRG | 242 | | |
| QY | 246 | PDETDELHYFNVTMTHYLNNGATRDKLXGVGFYGRAMSIDRSKLKLGDPKAGMSP | 305 | | |
| DB | 243 | SADTTDRLQINNVNASIHFWLAQGGCTGRKVLGIPLYGRNFTLAGAANTQIGAPTGGGT | 302 | | |
| QY | 306 | PGFISGEQVLSYIELCQLFQKEEWHIQYDEYNNAPYGNDKIWTGYDDLASISCLAPL | 365 | | |
| DB | 303 | VORYTREPQWGMNGYFCEKLAATEANDLWSEBQQPYAVRNQWVGYYDLRSVQLKVKYL | 362 | | |
| QY | 366 | KELGVGVVWVWLENDNDFKHC-GPKNPLNKNVHNMINGDEKNSFECILGGSTTPTPTT | 424 | | |
| DB | 363 | LQOGLGGMVWVWLENTDDPLGVCGGGRYPLMHHEPILNVNGGT-----PSTTMPSPSV | 413 | | |
| QY | 425 | TPTT | 484 | | |

47K glycoprotein p
hypothetical prote
chitinase (EC 3.2.)
chitinase (EC 3.2.)
chitinase (EC 3.2.)
chitinase (EC 3.2.)
chitinase (EC 3.2.)
hypothetical prote
chitinase - Strept
chitinase (EC 3.2.)
chitinase homolog
probable chitinase
probable membrane
chitinase homolog
chitinase chi-A or
chitinase - Autogr
chitinase (EC 3.2.)

Db 414 APTT-----SIVAPGTTTTTGANPGTTQPPPT--SDAPNHTTTTTEGNGFTTRPPSG 466
QY 485 TTTEHTSETPKYTTTVDGHLKCY--KEGDIPHPTNHHKYLVC-----BFVNGGWVWH 535
Db 467 -----DG-----PCAGGRYGFVPHPTNCARYVICLTADTYVEFT----- 500
QY 536 IMPCPPPGTIW 545
Db 501 ---CPFGILF 507
RESULT 2
T14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14075
R:de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A:Reference number: Z17872
A:Accession: T14075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1635
A:Cross-references: EMBL:AF026492; NID:G2564720; PID:G2564721; PIDN:AA81850.1
C:Genetics:
A:Gene: CHT2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 28.0%; Score 869; DB 2; Length 1635;
Best Local Similarity 34.3%; Pred. No. 4.8e-46;
Matches 196; Conservative 95; Mismatches 196; Indels 84; Gaps 18;
QY 21 IKRDNDYKSNM-----RIVCYGTWSVHKVD-PYTIIDIPFKTHLMYGFALIDE 73
Db 650 INKETSNDENAIESDVIDYKVCYFTNMAWYRQNGKYLPLDIDADLCTHIVYGFVILDR 709
QY 74 KYTTIQVDPYQDDNHSN---EKGYERFNNLRKLNPELMTWISLGGWYEGS-EKYSDM 129
Db 710 DRLVIK-----PHDSNADIDNRFYRVVEYKKGKKVT--VAIGWNDSAGKYSRL 759
QY 130 AANPTYROOFIQSVLDFQYKFDGLDLDWEYVPSRLGNPK-----IDKONYLALVRELKD 185
Db 760 VPSAARQKFIADVAFTIEKYFGDGLDLDWEYVPCVQVDCCKGFSDEKEGASLVVELSQ 819
QY 186 AFEPHYLLTAASPGKDKIDRAYDIKELNKFDMWNTYDYGWENFYGHNAPLYKR 245
Db 820 AFKPKGLLSAVSPSKVVDGYDVTLSDYMDIAVWAYDYHQWQDKTGHVAPMYEH 879
QY 246 PDETDELHTYFNNTMYLNGATRDKLUNGVPFYGRAMSIEDRSKLKLDPAKGMSP 305
Db 880 PDDFDKT---FNAFTIHYWIEKGDPRKLXMGMPYQGSFSLADNKEHGLNATYGGGE 936
QY 306 PGFISGEGLVSVIELCQLFOKEEWHIQVEYNN-APYGYNDKIWVGVDLLASISCKLAF 364
Db 937 AGESTRARGLSYEILCANIRKKMTVARDKRGMGPYAKGDQVQVDDQWIMRHKEY 996
QY 365 LKELGVGVMWWSLENDPFKHCQ-PKQPLLNKNVNMINGDEKNSFECIL--GPSSTTPTT 421
Db 997 VKAMGLGGAIIWALDLDLDFRNLCDCEYPLLATINRVLNYPGPRCVLEKEPQREPER 1056
QY 422 PTTTPT 481
Db 1057 PTPRPT 1104
QY 482 PPTTTEHTSETPKY-----TTVVDGHLIKCYKEGDIHPHTNHHKYLVCFFVNGW 532
Db 1105 PTTTTT---NVPSYQEIANEVDEPTCTDGLF-----VPHPTDCNKYICQYK--- 1150
QY 533 WWHIMPCPPGTIW-----CQEKLT 551

Db 1151 -----LCPGGLYWSVDHCDWPQSTNCRNKQT 1176
RESULT 3
A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C>Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C:Accession: A56596
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinase
A:Reference number: A56596; MUID:93357793; PMID:8353525
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:U02270; GB:S64757; NID:G406048; PID:G406049
A:Experimental source: larvae
A>Note: sequence extracted from NCBI backbone (NCBI:136417, NCBI:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
Query Match 27.0%; Score 838.5; DB 2; Length 554;
Best Local Similarity 35.4%; Pred. No. 9.2e-45;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;
QY 6 AILSIMACIGLMNASIKRDNDYKSNMIRIVCYGTWSVYHK-VDPYTTIEDIDPFKCTHL 64
Db 3 ATLATLAVLATAV-----QSDSRARIVCFVSNWAVRPGVGRYGIEDIPVEKCTHI 55
QY 65 MYGFAKIDYKTYIQVDFPYQDDNHSNWKRGYERFNNLRKLNPELMTWISLGGWYEGS 124
Db 56 IYFVIGVTEGSEVLIDPELD-----VDKNGFRNFTSLRSSHPSVKFVWVGWAGSGS 110
QY 125 KYSDMAANPYROOFIQSVLDFQYKFDGLDLDWEYVPSR-LGNPKIDKQNYLALVREL 183
Db 111 KYSHWVAQKSTRMSFRSVVSVFLKKYDFDGLDLDWEYVPSRSGGFSKDKPLYLQVEL 170
QY 184 KDAF--EPHYLLTAASPGKDKIDRAYDIKELNKFDMWNTYDYGWENFYGHNAPLY 241
Db 171 RRAFIRVGKWEITAAVPLANFELMEGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYRPPDETDELHTYFNNTMYLNGATRDKLUNGVPFYGRAMSIED----- 290
Db 231 LYRPHD-QWAYEKLNVNDGLHLWEKGCPSNKLVVGFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLKLDPKAGMSPGFGISGEGLVSVIELCQLFOKEE--WHIQVEYNNAPYGYNDKI 348
Db 290 NKEAGGDDPAPYTNATGF-----WAYEICTEVDKDDSGWTKQWDEQKCFYAYKGTQ 342
QY 349 WVGVDLLASISCKLAFKELGVGVMWWSLENDPFKHCQPKNPLNKNVNMINGDEKNS 408
Db 343 WVGYEDPSRVEIKKQNIKKQYLGAMTWALDMDDFQGLGCKGNPLIKLHKMS----- 396
QY 409 FECILGPSTTTPT 456
Db 397 -----SYTVPPPHENTTPTPEWARPSTPSDPSGDPITPTTTAKASITKTIVKIT 449
QY 457 TPTTPT 505
Db 450 TTTTAKPPQVIDEENDINVRPEPKPEPEPEVEFP-PTENE-----VDGSEI 497
QY 506 KCYKEGD-IHPHTNHHKYLVCFFVNGWVWHIMPCPPGIWCOEKLTC 552
Db 498 -CNSDQDIIPDKKCHKYWRG--VNGE--AMQFSQHGTVFVNVELNVC 540
RESULT 4
A53918
chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C>Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000

[illegible]

| | | | |
|---|-----|---|-----|
| QY | 188 | EPHG-----YLLTAAVSPGKIDIRAYDIKELNKLFDMMATWMTYDYHGCHENFYGHNA | 244 |
| DB | 165 | KNEAQLTWPRPLLISAAVSGDPHVQVQYKARLLGRLLDFISVLSDYDLHGSWEKVTGHS | 224 |
| QY | 241 | PLYKRPDETDELHTYFNNVNYTHYYLNNAGTRDKLVMGVPFYGRAWSIEDRSKILGLGDP | 300 |
| DB | 225 | PLFSLPGDPK-----SSAYAMNYWRQLGVPPPEKLLNGLTYGRTPHLLKASQNELRAQA | 278 |
| QY | 301 | KGMSPPGISBEGVLSYIELCOLFQKEEMHIQDEYNNAPYGYNDKIWGYDDDLASISC | 360 |
| DB | 279 | VGPASPGKYTKQAGFLAYEICCFVRRAKRWINDQV--VPYAFKGKEWGYDDAISFGY | 336 |
| QY | 361 | KLAFKLKELGVSGVWVNSLENDDEKGH-CGP-KNPLLNKVNMMINGDEKKSFECLIGPSTT | 418 |
| DB | 337 | KAFFIKREHFGGAWMTLUDLDDFGYFCGTGPPFLVHTLNNLNVNDEFSS----- | 386 |
| QY | 419 | TTPT-----TTTPTTTPTPTPTPTPTPTPTPT-----TTTPTPTPTPTPTPTPT | 466 |
| DB | 387 | TSPKFWSTAVNSSRIGPENPTWTRDLTTGLGLPPGGEAVATEHRSKSETWTITPKGE | 446 |
| QY | 467 | TTTTPT-TP-----TPAPTTTPSP-----TTTHTSETP | 494 |
| DB | 447 | IATPTRTPLSGRHTAAPEGKTESPGKPLTTVGHVAVSP | 466 |
| RESULT 7 | | | |
| I46470 | | | |
| estrogen dependent oviduct protein precursor - sheep | | | |
| C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep) | | | |
| C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999 | | | |
| C/Accession: I46470 | | | |
| R:DeSouza, M.M.; Murray, M.K. | | | |
| A:Title: An estrogen-dependent secretory protein, which shares identity with ch | | | |
| and embryo development. | | | |
| A:Reference number: I46470; MUID:95269691; PMID:7750470 | | | |
| A/Accession: I46470 | | | |
| A>Status: preliminary; translated from GB/EMBL/DBJ | | | |
| A:Molecule type: mRNA | | | |
| A:Residues: 1-539 <DES> | | | |
| A:Cross-references: EMBL:U16719; NID:9885600; PIDN:AAC48471.1; PID:9885601 | | | |
| Query Match 22.0%; Score 685; DB 2; Length 539; | | | |
| Best Local Similarity 33.8%; Pred. NO. 3.2e-35; | | | |
| Matches 173; Conservative 90; Mismatches 185; Indels 64; Gaps 17 | | | |
| QY | 10 | IVACIGLMAVASIKRDNDYSKNPMRIIVCYGTWSVYHKVDPIYI--EDIDPFKCTHLMYG | 67 |
| DB | 4 | LLLVGLL---LMLKHGDAH--KLVCYFTNWA-FSRGSASLPRDLDPFLCTHLVFA | 57 |
| QY | 68 | FAKIDYKYTIQVDFPYODDHNHNSWEKRGYERFNNLAKNPELTTMISLGGWYEGSEKYS | 127 |
| DB | 58 | PASMNNOQ--IVPKDPLD-----EKLIYEFNKLKERNRGLKTLISVGGWVFGTSRT | 108 |
| QY | 128 | DMAANPTYRQQTQSVLDFQAEYKPDGLDLDWYVPSGRSLGNPKIDKONYALVRELKDAF | 187 |
| DB | 109 | KMLSTFSNRERFVKSVIALLRTHGPDGLDFFLYPGLR-GSPARDRWTFVFLLEELQAF | 167 |
| QY | 188 | EPHG-----YLLTAAVSPGKIDIRAYDIKELNKLFDMMATWMTYDYHGCHENFYGHNA | 244 |
| DB | 168 | KNEAQLTWPRPLLISAAVSGDPHVQVQYKARLLGRLLDFISVLSDYDLHGSWEKVTGHS | 227 |
| QY | 241 | PLYKRPDETDELHTYFNNVNYTHYYLNNAGTRDKLVMGVPFYGRAWSIEDRSKILGLGDP | 300 |
| DB | 228 | PLFSLPGDPK-----SSAYAMNYWRQLGVPPPEKLLNGLTYGRTPHLLKASQNELCAQA | 281 |
| QY | 301 | KGMSPPGISBEGVLSYIELCOLFQKEEMHIQDEYNNAPYGYNDKIWGYDDDLASISC | 360 |
| DB | 282 | AGPASPGKYTKQAGFLAYEYVCSFQRAKRWINDQV--VPYAFKGKEWGYDDAISFGY | 339 |
| QY | 361 | KLAFKLKELGVSGVWVNSLENDDEKGH-CGP-KNPLLNKVNMMINGDEKKSFECLIGPSTT | 418 |
| DB | 340 | KAFFIKREHFGGAWMTLUDLDDFRNGFCGTGPPFLAHTLNNLNVNDEFSS----- | 389 |

```

QY 419 TTPP-----TTTTTTTTTTTSPSPPTTPTTPTPTP-----TTTTPTPTTPTTPTSPPT 466
DB 390 TSPSPFWFSTAYNSRIGPEMPTTRDLTGTGLPLGGEAVATEPHRKSATMTTTPRGE 449
QY 467 TTPPTP-TP-----TPAPTSTPSPTTTEHTS 491
DB 450 TATPTNTPLUSSGRPTAAPGKTESPGEKPLTS 481

RESULT 8
I38605
oviductal glycoprotein - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I38605
R:Arias, E.B.; Verhage, H.G.; Jaffe, R.C.
Biol. Reprod. 51, 685-694, 1994
A:Title: Complementary deoxyribonucleic acid cloning and molecular characterization
A:Reference number: I38605; WUID:95119256; PMID:7819450
A:Accession: I38605
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-654 <RES>
A:Cross-references: EMBL:U09550; NID:G529147; PID:G529148

Query Match 21.5%; Score 668; DB 2; Length 654;
Best Local Similarity 33.1%; Pred. No. 4.7e-34;
Matches 176; Conservative 91; Mismatches 192; Indels 72; Gaps 19;

QY 10 IMACIGLGNASIKRDNHYSKNPMRIVCVGVTSVYKVDPYTI--EDIDPFKCTHLMYG 67
DB 4 LLLWGLV---LVLKHDGAH--KLVCYFTNA--HSRPGPASILPHDLPFLCTHLIPA 57
QY 68 FAKIDYKYTIQVDFPYQDNNHNSWKEGVERFNNRLKNPELTTWISLGGWVEGSEKYS 127
DB 58 FASMNNOI---VAKDLQD-----EKILYPEFNKLKERNRELKTLISIGWNFGTSRFT 108
QY 128 DMAANTYQOFTQSVLDFLOEYKFGDLDDWEPGSRCLGNPKIDKQNYLALVRELKDAF 187
DB 109 TWLSTFANREKFTIASVISLRTDFDGLDLFLYPGLR--GSPMHDRTWTFLLIEELLFAP 167
QY 188 EPHG-----YLLTAASVSGDKIDRAYDIKELNKLFDMMVNTYDYHGWENFYGHNA 240
DB 168 RKEALLTMPRLLSAAVSGVPHIVQTSYDVRFLGRLDLPINVLSDLHGSWERFTGHS 227
QY 241 PLYKRDEDELHTYFNFNVTMYHYLLNNGATDKLVGVVFFYGRAWSIEDRSKLKLGDP 300
DB 228 PLFSLFEDPK-----SSAVANNYWRKLGAPSEKLINGIFTYGRTELLKASKNGLQARA 281
QY 301 KMSPPGFTSGEGVLSYELCOLF--QKEWHIQYDEYNAPYGNDKTIWGVYDDLASI 358
DB 282 IGPASPGKVTQEGFLAYFEICSFVNGAKGHV---IDYQVPEYANKGEWVGYDNAISF 337
QY 359 SKLALFKELGVSGVMVWSLENDDFKG--HCGP--KNPLLKNVHNMINGDEKNSFECILGPS 416
DB 338 SYKAWPIRREHFHFGAMVWTLDMDDVDRGTFCGTGFFPLVYVINDILVRAEFS-----S 389
QY 417 TTTPTP--PTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 466
DB 390 TSLPQWLSSAVNSSDTERLATVTTAWTDSKILPPGEGAGVTEHGKCNWITITERGT 449
QY 467 TPTPTTPT-----PAPTSTPSPTTTEHTSETP--KYTYVVDGH 503
DB 450 TVTPTKETVSLKGHTVALGEKTEITGAMTWTSVGHQSWTFGEKALTPV--GH 499

RESULT 9
A49562
cartilage glycoprotein gp39 precursor - human
N:Alternate names: 39K synovial protein
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999

```

C;Accession: A49562; S10677; A33162
R;Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A;Title: Human cartilage gp-39, a major secretory product of articular chondrocytes and
A;Reference number: A49562; MUID:94064658; PMID:8245017
A;Accession: A49562
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAK>
A;Cross-references: GB:M80927; NID:g348911; PIDN:AAAL6074.1; PID:g348912
Biochem. J. 269, 265-268, 1990
A;Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary proteo
A;Reference number: S10677; MUID:90328983; PMID:2375755
A;Accession: S10677
A;Molecule type: protein
A;Residues: 22-40,'X',42-45 <NY2>
C;Superfamily: Streptomyces chitinase chi40
C;Keywords: cartilage; extracellular protein; glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 21.3%; Score 662; DB 2; Length 383;
Best Local Similarity 34.9%; Pred. No. 5.5e-34;
Matches 144; Conservative 81; Mismatches 136; Indels 52; Gaps 12;

Qy 2 KTIYAILISIMACIGLMNASIKRDNDYSKNPMPRIYCVGTWSVYHKVDPTYIED-IDPFK 60
:
Db 7 QTGFVLVLLQQ-----SAYKLVCYTSSOYREGDSCFPDALDRFL 50

Qy 61 CTHLWYGFAKIDEKYITQVDPDYDDNHNSWEKG---YERFNULRNKPFLTWTSLG 117
:
Db 51 CTHIIVSFANI-----SNHDITWEWNVDVLYGMENLTNKRNFNLKTLISVG 97

Qy 118 GWTEGSSEKYSDMAANPYROFTQSULDPLQEYKFQGLDLWDYPSRLGNPKDKONYL 177
:
Db 98 GWNPGSORFKIASNTOSRETFIKSVPPFLRTHTFGDGLAWLYPCRR-----DKQHFT 151

Qy 178 ALVRELKDAF-----EP--HGVLITAAVSCKDIKDRAVDIKELKLFDMVNVTYDVHG 231
:
Db 152 TLTKEMAEFIKEAQPKQLLSAUSAGKVTDISSYDAKTSQMLDFISINTYDFHGA 211

Qy 232 WENFYGHNAPLYKRPDETBELHYFNYNVNMHYLNNGATRDCLVMGVPPFYGRAWSIEDR 291
:
Db 212 WRGTTGHSPLF-RGQEDASPDSPSTNDYAVGYMLRALGAPASKLVMGITPTFGSFTLAS- 269

Qy 292 SKULGDPAKMSPGPFTSGEGLSYIEICQLFQKEWHIQYDEVYNAPYGVNDKIWVG 351
:
Db 270 SETGVGAPISGPGIPGRFTKEAGTAYEICDFLRGATVHRTLQG--QVPYATKGNQWVG 327

Qy 352 YDDLASISCKLAFLKELGVSGVMWSLENDDFKG-HCGP--KNPLLKNVHMNI 401
:
Db 328 YDDQESYKSVQYLKDRQLAGAMVWALDLLDFQSGFGQDLRFPLTNKDAL 380

RESULT 10
S51327
heparin-binding glycoprotein 38k - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
C;Accession: S51327
R;Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A;Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in differ
A;Reference number: S51327
A;Accession: S51327
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <SHA>
A;Cross-references: EMBL:D47803; NID:g634097; PIDN:CRAA87764.1; PID:g634098
C;Superfamily: Streptomyces chitinase chi40

Query Match 20.9%; Score 650; DB 2; Length 383;

```
QY 150 YKFDGLDLDWYPSGRLGNPKIDKONYLALVRELKADP-----EP--HGYYLLTAASPGKD 203
D 131 YGFDGLDLAWLPPRLR-----DKQYFTLLKELNAETKEVQPGREKLLLSAALSAGKV 184
QY 204 KIDRAYDTELKELNKLDPWNVVNTYDYGWENFYGHNAPLYKRPDETDELHYTFNVNTYTMH 263
D 185 AIDTGYDIAQAOLHDFNLMTYDFHGVWROITGHHSPLFGQKDT-RFRYSNVNTYAVQ 243
QY 264 YLLNNGATRDILVMGVPYPGRAWSTEDRSKLGDPKAGMPPGFIAGEGVLSYIELCQ 323
D 244 YNTRIGAASKLLMGIPFGKSFILAS-SENQLGAPTISGEGLPORFKEAGTLAYEILD 302
QY 324 LFOKEEMHIQDYENYAPYNDKIWGYDDLASISCKLAFELKELGVSQV-----MWSLE 379
D 303 FLKGAEVHRLSNE--KVFPATKQWGVYEDKESVKNKVGFLKELKLAGAWCGHWINWI- 359
QY 380 NDDFKGCHGCPKNPLLNKVNHNMGDKNSFECILGPTTTTPTTTP 426
D 360 ----SGHCQPKN-----SSRSTPSRMP 378

RESULT 12
S27879
secretory protein YM-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999
C:Accession: S27879
R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.
submitted to the EMBL Data library, June 1992
A:Description: Molecular characterization of a secretory protein (YM-1) transiently exp
A:Reference number: S27879
A:Accession: S27879
A:Molecule type: mRNA
A:Residues: 1-399 <CHA>
A:Cross-references: EMBL:M94584; NID:G202441; PIDN:AAB62394.1; PID:G202442
C:Superfamily: Streptomyces chitinase chi40
F.1-21/Domain: signal sequence #status predicted <SIG>
F.122-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 19.6%; Score 608; DB 2; Length 399;
Best Local Similarity 36.3%; Pred. No. 1.3e-30;
Matches 141; Conservative 67; Mismatches 140; Indels 40; Gaps 14;

QY 34 RIVCVGTWVYHKVD-PYTTIEDIDPFKCTHLMYGFAKI--DEKYIIQVDFDPYQDDNHN 90
D 23 QLMCYTTSWAKDRPIEGSKFGKGNIDPCLCITLIYAFAGMGNNEITYT-----H 70
QY 91 SWEKRGYERFNLR-LKNPELTMTISLGWYEGSEKYSDMAANPYRQQTQSVDLFLQE 149
D 71 EODLRDYALNGLKDKKNTLKLTAIGWKFPGAPPSAMVSTPQNROIFQSVIRFLRQ 130
QY 150 YKFDGLDLDWYPSGRLGNPKIDKONYLALVRELKADPEPHG-----YLLTAASPGK 202
D 131 YNFDGLNLDWQYPSGR-GSPPKDHLFSLVLYKMKAFEEESVEKIDPRLLLT---STGA 186
QY 203 DKIDRAYD-IKELN--KLDFMNVNTYDYGWENFYGHNAPLYKRPDETDELHYTFNVN 259
D 187 GIIDVIKSGTRSLNCLSLLDYIQVNTYDHPDKDGTGNSPLYKSPYDICK-SADLNDV 245
QY 260 YMHYLLNNGATRDILVMGVPYPGRAWSTEDRSKLGDPKAGMPPGFIAGEGVLSYI 319
D 246 SLISYWKDHGAASEKLIIVGFPAHYGTFTILSDPSKTGIGAPTISTGPPCKYDDESGLLAY 305
QY 320 ELQCLPQK---BEWHIQDYENYAPYNDKIWGYDDLASISCKLAFELKELGVSQVNV 376
D 306 EVCYTFNEGATEWV----DAQEVFYAQGNWGVYDNVRGFKUKAQWLKONNLGGAVVW 361
QY 377 SLENDDFKXG-HCGPKN-PLLNKVNHNIN 402
D 362 PLDMDDFSGSFCHQRHPTLTSTLKGLDN 389

RESULT 13
```

```
A38221
chitinase (EC 3.2.1.14) MF1 - nematode (Brugia malayi)
C:Species: Brugia malayi
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A38221
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Piessens, W.F.; Perlier, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmission-blocking antibodies recognize microfilarial chitinase in brugian l
A:Reference number: A38221; MUID:92179220; PMID:1542646
A:Accession: A38221
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <FUH>
A:Cross-references: GB:M73689; NID:G156063; PIDN:AAA27854.1; PID:G156064
A:Note: sequence extracted from NCBI backbone (NCBIP:85345)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 19.5%; Score 604.5; DB 2; Length 504;
Best Local Similarity 30.1%; Pred. No. 3e-30;
Matches 156; Conservative 86; Mismatches 208; Indels 69; Gaps 17;

QY 37 CYVGTWVYHKVD-PYTTIEDIDPFKCTHLMYGFAKIDYKTIQVDFPYQ-DDNHNSEK 94
D 27 CYTWNQAQYRDEGKFLPCNIFNGLCTHLYAFKVDK-----LGSKPFWEHDETEWSK 82
QY 95 RGYERFNRLKKNPELTMTISLGWYEGSEKYSDMAANPYRQQTQSVDLFLQEKYKPDG 154
D 83 GMSAVTKURETNPGKLVLLSYGVNFGSAITFTAKSAQKTERPIKSAIAFLRKNPFG 142
QY 155 LDLDWEYP-GSRLGNPKIDKONYLALVRELKADPEPHGYLLTAASVSPGKDKIDRAYDIKE 213
D 143 FDLDEYFVGVAAEHAKEAMKTAPEAEAKTSGKOR-LLLTAASVAGKTIDGSVNVES 201
QY 214 LNLKFMNVNMYDYHGGWENFYGHNAPLYKRPDETDELHYTFNVNTYMHYLLNNGATRD 273
D 202 LKGNFOLLFLMSYDLHSGWENKVDLHGLHPTKGEVSGI-GIENTEFAADYVASKMPKE 260
QY 274 KLVMGVPYFGRAWSTEDRSKLGDPKAGMSPGPFISGEGVLSYIELCOLFKQSEWHIQ 333
D 261 KIITGIPMAQGWTLDPNSETAIGAAASPSASKNTPAGGTASWEICKYKKEGKETV 320
QY 334 YDYVYNAPYGYNDKIWGYDDLASISCKLAFELKELGVSQVWMSLENDDFPKG-HCGP-KN 391
D 321 HQEGVGA-YWVGQDQWGYDNEETIRIKMKWLKEGYGGAFIWDLDFDFTGKSCGKGY 379
QY 392 PLLNKVNHNMGDKNSFECILGPTTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 451
D 380 PLINAISSLEGESEN-----PEITTEPSITTEAVETDETEB----- 418
QY 452 TTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 507
D 419 -----TSETEAYDTD-----ETEETSET-EATYTDTETEG--EC 451
QY 508 -YKEGDIHPHTNTHKYLVCVFNGVWVHIMPCPPGTIW 545
D 452 PERDGLFPHTDCHLFQIC-----ANNIAYVMQCPATFF 486

RESULT 14
D38764
chitinase BHO916 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D38764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D38764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-599 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:G10173440; PIDN:BA04635.1; GSFDB:GN001
```

A; Experimental source: strain C-125

C; Genetics:

A; Gene: BH0916

Query Match 19.1%; Score 592; DB 2; Length 599;
Best Local Similarity 29.0%; Pred. No. 2.3e-29;
Matches 183; Conservative 98; Mismatches 216; Indels 134; Gaps 29;

```
QY 2 KTIYAILSMACIGLMNASIKRHDNDYKSNRIVCVGTWTSVYHKVDYTTIEDIDPFKC 61
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 KLLIWIWTVLLVWTPADQVGAETSDDQYKIVAYPSWAGYGR--DYQVMDIDASKI 66
QY 62 THLMYGFAK-----IDYKVTIQVDFPYQD-----DNH 89
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 SHNYAFANICWGRHGNPDAGNPQTWSQDENGVIDVPNGSIYMGDFWIDAQKSNP 126
QY 90 NSWEK--RG-YERFNNLRKLNKPNLTWISLGGHYEGSEKYSMAANPTVRQFIOQLV 146
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 DTWDEPLRGFKQINKLKEHPHLKLLISVGGH-TWSNRFSDWAATKETRENFANSA 185
QY 147 LOBYKFDGLDLDWEYP--GSRLLGNPK--IDKQNYLALVRELKDAEPHG----- 196
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 IRKYGFQVDVDWEYFVSGGLPGNSRRPDKENHVLILQEVRLKDEAGQEDGKYLL 245
QY 197 A--VSPGKOKIDRAYDIKELNKLFDWNNVTYDHGWNENFYGHNAPLYKRPDETD- 253
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 ASGASPGYVENNK---LNETAETVDNINIMTYDFNGWQWISGHNAPLYDPATAN 302
QY 254 T--YENVNVTMYYLNNGATRDKLVMGVPEYGRWSTEDRSKLGDPKAGMSPPGF 311
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 TPSEHFNVESAVEGHLQAGVFEHKLVLGMPPYGRGWSNCDGA--NQGEYQR-C 359
QY 312 EEGVLSVIELCOLFQKEWHIQ--YDEYN--APYGN--DKIWWGVDDILASISCKL 362
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 ENGUFDFSDL-----EDHVNKNGQRYNDVAKVPFLYNATNGNFIYDDEES 413
QY 363 AFLKELGVSGVMVWSLENDDFKHCGKPKNLLKNVNMINGDEKNSFECILGPST 422
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 DFIKSNLAGSMFWDVSGDR-----NGTLTALADQLG---FTPHE 451
QY 423 TTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 476
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 GQPEEPSAPTNIQATEVISTVTLTWAQPTBEPQSVAYDSKEKTTHTTTIEDLQ 511
QY 477 PTTSTPSTPTTEH-----TSETPKYTYVDG-----HLIKCYKEGD-IP 519
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 512 PETTYFVVSAEKHQGRHAGALQVTTKSETGGDGTAPTQWANNVYTGQVQHGG 571
QY 520 HKYLVCBFVNGWVWVHMPCPGT-----IW 545
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 572 YE-----AKWWT--TGEEPGTTGEWGVW 592
```

RESULT 15

A38368

Chitinase (EC 3.2.1.14) precursor - Bacillus circulans

C; Species: Bacillus circulans

C; Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 15-Oct-1999

C; Accession: A38368

R; Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.

J. Biol. Chem. 265, 15659-15665, 1990

A; Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolution

A; Reference number: A38368; MUID: 90368776; PMID: 2203782

A; Accession: A38368

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-699 <WAT>

A; Cross-references: GB:M57601; GB:J05599; NID:gl066341; PIDN:AAA81528.1; PID:gl142688

C; Superfamily: fibronectin type III repeat homology

C; Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match

Best Local Similarity 18.1%; Score 561; DB 2; Length 699;

29.4%; Pred. No. 2.3e-27;

```
Matches 170; Conservative 67; Mismatches 185; Indels 156; Gaps 23;
QY 34 RIVCVGTWTSVYHKVDYTTIEDIDPFKC THLMYGFAKI----- 71
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 KIVGYPSWAAYGR--NYNVADIDPTKVTHTINYAFADICWNGIHGNDPSPGNPVTWTCQ 102
QY 72 DEYKVTIQV-----PDPYQDDNH---NSWEK--RGYERFNNLRKLNKPNLTWISL 117
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 NEKQTTINPNGTIVLGDPIWDTGKTFTAGDTWDOPIAGNINQLNKLKQTNPNLTIIS 162
QY 118 GWYEGSEKYSMAANPTVRQFIOQLVDFLOEYKFDGLDLDWEYF--GSRLLGNPK-- 173
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 GW-TWSNRFSDWAATREVFANSAVDFLRKYNFDGVDLDWEYFVSGGLDGNKSRPED 221
QY 174 QNYLALVRELKDAEPHG-----YLLTAAVSPGKOKIDRAYDIKELNKLFDWNNVT 227
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 QNYTLLSKIREKLDAAGAVDGKYLTLIA--SGASATYAANTELAATAAIVDWINIMTY 280
QY 228 YHGWNENFYGHNAPLYKRPDET---DELHTYFNNVTMYYLNNGATRDKLVMGVPE 283
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 FNGAWQKISAHNAFLNVDPAASAAGVDPDANT--FNVAAGAQLHLDAGVPAAKLV 339
QY 284 RAWSTIEDRSKLGDPKAGMSPPGFISGEGLVSVIELCOLFQKEWHIQ---YDEYN- 339
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 RWDGCAQAGNGQYQTCGGSSVG--TWEAGSFDFYDL-----EANYINKNGYTRW 391
QY 340 ---APYGN--DKIWWGVDDILASISCKLAFKELGVSGVMVWSLENDDFKHCGK 394
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 392 TAKVPYLYNASNKRPFISYDDAESVGKYTAIYKSKGLGAMFWELSGDRNK----- 441
QY 395 NKVNMINGDEKNSFECILGPSTTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 436
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 -TLQNKLKAD-----LPTGGTVFVDDTTAPSVPGNARSTGVTANSVTLAWN 486
QY 437 -----SPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 458
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 ASTDNVGVTVGNVNGANLATSVTGTTATISGLTAGTSYFTTIKAKDAAGNLSA 546
QY 459 PTTTPTSP--TTPTPTPTPTP--APTSTPTPTPTTTEHTSET 493
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 547 VSTTAQPGDQAPTAPTNLASTAQTTSITLSWTAST 584
```

Search completed: March 22, 2004, 07:01:28

Job time : 37.0951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:31:13 ; Search time 21.2868 Seconds
(without alignments)
1357.597 Million cell updates/sec

Title: US-09-662-293-15

Perfect score: 3107

Sequence: 1 MKTIYAILSLIMACIGLMNAS.....IMPCPPGTIWCQKLTGICE 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 838.5 | 27.0 | 554 | 1 | CHIT MANSE |
| 2 | 738 | 23.8 | 465 | 1 | CHT1_HUMAN |
| 3 | 734.5 | 23.6 | 473 | 1 | CHIA_MOUSE |
| 4 | 723 | 23.3 | 508 | 1 | CHIL_DROME |
| 5 | 708.5 | 22.8 | 476 | 1 | CHIA_HUMAN |
| 6 | 705.5 | 22.7 | 617 | 1 | CHIT_CABEL |
| 7 | 696.5 | 22.4 | 527 | 1 | OGP_FIG |
| 8 | 685.5 | 22.1 | 537 | 1 | OGP_BOVIN |
| 9 | 685 | 22.0 | 539 | 1 | OGP_SHEEP |
| 10 | 680 | 21.9 | 381 | 1 | C3L1_MOUSE |
| 11 | 680 | 21.9 | 671 | 1 | OGP_MESAU |
| 12 | 680 | 21.9 | 721 | 1 | OGP_MOUSE |
| 13 | 668 | 21.5 | 678 | 1 | OGP_HUMAN |
| 14 | 662 | 21.3 | 383 | 1 | C3L1_HUMAN |
| 15 | 654.5 | 21.1 | 623 | 1 | OGP_PAPAN |
| 16 | 630 | 20.3 | 390 | 1 | C3L2_HUMAN |
| 17 | 611.5 | 19.7 | 396 | 1 | CHT1_MOUSE |
| 18 | 604.5 | 19.5 | 504 | 1 | CHIT_BRUMA |
| 19 | 576 | 18.5 | 458 | 1 | CH13_DROME |
| 20 | 561 | 18.1 | 699 | 1 | CH11_BACCI |
| 21 | 411 | 13.2 | 423 | 1 | CH14_TRIHA |
| 22 | 410.5 | 13.2 | 423 | 1 | CH11_APHAL |
| 23 | 407.5 | 13.1 | 563 | 1 | CHIA_SERMA |
| 24 | 384 | 12.4 | 820 | 1 | CHIA_ALTSO |
| 25 | 377.5 | 12.1 | 550 | 1 | CHIT_NPVOPL |
| 26 | 360.5 | 11.6 | 551 | 1 | CHIT_NPVAC |
| 27 | 353 | 11.4 | 499 | 1 | CHIB_SERMA |
| 28 | 341 | 11.0 | 619 | 1 | CHIT_STRLI |
| 29 | 339.5 | 10.9 | 427 | 1 | CH11_COCPO |
| 30 | 326 | 10.5 | 610 | 1 | CHIT_STREPL |
| 31 | 280 | 9.0 | 1046 | 1 | CH1D_VIBFU |
| 32 | 266.5 | 8.6 | 474 | 1 | VTP3_TTVIV |
| 33 | 259 | 8.3 | 5179 | 1 | MUC2_HUMAN |

ALIGNMENTS

RESULT 1

CHIT MANSE

ID CHIT MANSE STANDARD; PRT; 554 AA.
AC P36362;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Endochitinase precursor (EC 3.2.1.14).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;
OC Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9335793; PubMed=8353525;
RA Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.;
RT "Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of Manduca sexta.";
RL Insect Biochem. Mol. Biol. 23:691-701(1993).
RN [2]

SEQUENCE FROM N.A.

RP MEDLINE=97215580; PubMed=9061927;
RX Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.;
RA "Isolation and characterization of a genomic clone for the gene of an insect molting enzyme, chitinase.";
RL Insect Biochem. Mol. Biol. 27:37-47(1997).
CC -!- FUNCTION: Digest chitin in the exoskeleton during the molting process.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Epidermis and gut.
CC -!- DEVELOPMENTAL STAGE: High levels seen in the epidermis on day 0, but rapidly disappears and is undetected on days 1-4 of fifth instar. It reappears on day 5 and peaks on day 7 after which a rapid decline is seen. In the gut is detected on day 6 with lower levels seen on days 0, 7 and 8.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl hydrolases).
CC -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U02270; AAC04924.1; -.
CC EMBL: L49234; AAB53952.1; -.
CC PIR: A56596; A56596.
CC InterPro: IPR002557; Chitin_bind_Pera.
CC InterPro: IPR001223; Glyco_hydro_18.
CC InterPro: IPR001579; Glyco_hydro_18AS.
CC -----

34 247 7.9 1513 1 MUC2_RAT
35 244 7.9 662 1 MUC1_XENLA
36 238 7.7 279 1 Y091_NPVOPL
37 230 7.4 400 1 MUA1_XENLA
38 227.5 7.3 1331 1 MANB_CALSA
39 221 7.1 267 1 EXTN_WAIZE
40 217.5 7.0 544 1 GP10_DICDI
41 217.5 7.0 1161 1 DANA_YEAST
42 215.5 6.9 705 1 GUN6_DICDI
43 215.5 6.9 1146 1 KTXA_KLULA
44 208.5 6.7 283 1 EXTN_SORBI
45 207 6.7 307 1 SGS3_DROME
Q62635 rattus norv
Q05049 xenopus lae
P10341 oryza psu
P10667 xenopus lae
P22533 caldocellum
P14918 zea mays m
Q08885 dictyosteli
P47179 saccharomyc
P22699 dictyosteli
P09805 kluyveromyc
P24152 sorghum bic
P02840 drosophila

```
DR Pfam; PF01607; CBM_14; 1.
DR DR PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR DR SMART; SM00636; Glyco_18; 1.
DR DR PROSITE; PS05040; CHIT_BIND_II; 1.
DR DR PROSITE; PS10095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 554 ENDOCHITININASE.
FT DOMAIN 396 453 SER/THR-RICH.
FT DOMAIN 495 553 CHITIN-BINDING TYPE-2.
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 554 AA; 62203 MW; 3989D756C96CD490 CRC64;

Query Match 27.0%; Score 838.5; DB 1; Length 554;
Best Local Similarity 35.4%; Pred. No. 6.3e-39;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;

QY 6 ALLSTIMACIGLWNASIKEDNDYDKNPRLVYVGTWGVYHK-VDPYTIEDIDPFKCTHL 64
Db 3 ATLATLAVLATAV-----QDSRRIRVYFSNWAAYRPGVGRYGIEDIPVEKCTHI 55
QY 65 MYGFAKIDEXKYTIQVDFPYODDNHNSWEKGYERFNNLRNKLNPETLTMISLGGWYGESE 124
Db 56 IYSFICVTEGSEVLIIDPELD-----VDKNGFRNFTSLRSHSPKRVKFWAVGWAAGSS 110
QY 125 KYSDMAANPTYRQFIQSLVDFLOEYKDFGLDWEYEGSR-LGNPKIDKQNYLALVREL 183
Db 111 KYSHVAAQKSTRMSFIRSVFSLKXYDFDGLDWEYFGAADRGGSFSDKDKLYLVQEL 170
QY 184 KDAP--BHGVLITAAVPGKDKIDRAYDIKELKLFDMWVWYDVGHWENYGHVAP 241
Db 171 RRAPIRVGKGWELTAAVPLANFRIMEGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDEIDELHTYFNVNTVHYLYNNGATDKLVMGVPFYGRAWSIED----- 290
Db 231 LYKRPD--QWYKELNVNDGLHLWEKGCPSNKLWVGIFYGRSFTLSAGNNNYGLGTFFI 289
QY 291 RSKLGLDPAPKMSPPGISBEGVLSYELCOLFOKEE--WHQYDYYNAPYGYNDKI 348
Db 290 NKEAGGPPAPYNTATGP-----WAYVEICTEVDKDDSGWKNDQGCPCYAYKGTQ 342
QY 349 WGYDDLASISCKLAFLKELGVSGVWVWSLENDPFKHCGRPNLLKLVKNNVINGDEKNS 408
Db 343 WGYEDPRSVETKMWIKQKQYLGAMTWAIDMDDFQGLCGEKNPLIKILRHMS----- 396
QY 409 FECILGPSTTTPTPTTTPTPT-----TPTTPPS-----PTPTTPSPSTTTPTTSP 456
Db 397 -----SYTVPPTPTPTPTPEWARPPSPSPSEGDPIPTTTAKPASTTKTKTKTT 449
QY 457 TPTPTTPSPST-----PTP--TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 505
Db 450 TTTTAKPQSVDENDINRPEKPEQPEVEVP-PTNE-----VDGSEI 497
QY 506 KYCKEGD-IPHTNIHKYLCVFVNGVWVHMPCPGPTIMCQEKLC 552
Db 498 -CNSQDQYIPDKKHCDKYRC--VNGE--AMQFSQHGTVFVNLNVC 540

RESULT 2
ID CHIT1 HUMAN STANDARD; PRT; 466 AA.
AC Q13231; Q9H3V8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chitotriosidase 1 precursor (EC 3.2.1.14) (Chitinase 1).
```

```
GN CHIT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND SUBCELLULAR
RP LOCATION.
RC TISSUE=Macrophage;
RX MEDLINE=96064695; PubMed=7592832;
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J.,
RA Aerts J.M.F.G.;
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT produced by macrophages";
RL J. Biol. Chem. 270:26252-26256(1995).
RN [2]
RP SEQUENCE OF 22-42 AND 178-198, FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=95138187; PubMed=7836450;
RA Renkema G.H., Boot R.G., Muijsers A.O., Donker-Koogman W.E.,
RA Aerts J.M.F.G.;
RT "Purification and characterization of human chitotriosidase, a novel
RT member of the chitinase family of proteins.";
RL J. Biol. Chem. 270:2198-2202(1995).
RN [3]
RP POLYMORPHISM, AND DESCRIPTION OF ISOFORM 3.
RX MEDLINE=98421482; PubMed=9748235;
RA Boot R.G., Renkema G.H., Verhoeck M., Strijland A., Bliet J.,
RA de Meuleneester T.M.A.M.O., Mannens M.M.A.M., Aerts J.M.F.G.;
RT "The human chitotriosidase gene. Nature of inherited enzyme
RT deficiency.";
RL J. Biol. Chem. 273:25680-25685(1998).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 22-386 IN COMPLEX WITH
RP CHITOBIOSE AND ALLOSAMIDIN.
RX MEDLINE=22095530; PubMed=11960986;
RA Fusetti F., von Moeller H., Houston D., Rozeboom H.J., Dijkstra B.W.,
RA Boot R.G., Aerts J.M.F.G., van Aalten D.M.F.;
RT "Structure of human chitotriosidase. Implications for specific
RT inhibitor design and function of mammalian chitinase-like lectins.";
RL J. Biol. Chem. 277:25537-25544(2002).
CC -!- FUNCTION: Degrades chitin and chitotriose. May participate in the
CC defense against nematodes and other pathogens. Isoform 3 has no
CC enzymatic activity.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted. A small proportion is lysosomal.
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1;
CC IsoId=Q13231-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q13231-2; Sequence=VSP_008631, VSP_008632;
CC Name=3;
CC IsoId=Q13231-3; Sequence=VSP_008633;
CC Note=Duplication of 24 bp in exon 10 leads to the use of a
CC cryptic splice site. The normal splice site is still present but
CC not used;
CC -!- TISSUE SPECIFICITY: Detected in spleen. Secreted by cultured
CC macrophages.
CC -!- POLYMORPHISM: A 24 bp duplication in exon 10 leads to the
CC activation of an alternative splice site and the production of an
CC inactive protein. About 6% of the population are deficient for
CC CHIT1 activity, while 35% are carriers and show reduced enzyme
CC levels. People with CHIT1 deficiency appear perfectly healthy.
CC -!- MISCELLANEOUS: Patients with type 1 Gaucher disease (GD 1)
CC [MIM:230800] have very high plasma levels of CHIT1, and this can
CC be used as diagnostic aid and to evaluate the success of
CC treatment. Successful therapy brings the CHIT1 activity levels
CC back to normal.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29615; AAC50246.1; --
CC EMBL; U62662; AAG10644.1; --
CC PDB; 1GUV; 31-JAN-02.
CC PDB; 1LG3; 18-SEP-02.
CC PDB; 1LG2; 18-SEP-02.
CC PDB; 1LQ0; 29-JUL-03.
CC Genew; HGNC:1936; CHITI.
CC MIM; 60031; --
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0004568; F:chitinase activity; TAS.
CC GO; GO:0009617; P:response to bacteria; TAS.
CC GO; GO:0009613; P:response to pest/pathogen/parasite; TAS.
CC InterPro; IPR002557; Chitin bind Peta.
CC InterPro; IPR001223; Glyco_hydro_18.
CC InterPro; IPR001579; Glyco_hydro_18AS.
CC Pfam; PF01607; CEM_14; 1.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC SMART; SM00494; ChtBD2; 1.
CC SMART; SM00636; Glyco_18; 1.
CC PROSITE; PS50940; CHIT_BIND_II; 1.
CC PROSITE; PS01095; CHITINASE_18; 1.
CC KEGG; Carbohydrate metabolism; Chitin degradation;
CC KEGG; Polysaccharide degradation; Hydrolase; Glycosidase; Chitin-binding;
CC KEGG; Signal; Alternative splicing; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 466 CHITOTRIOSIDASE 1.
FT DOMAIN 417 466 CHITIN-BINDING TYPE-2.
FT ACT SITE 140 140 BY SIMILARITY.
FT DISULFID 26 51
FT DISULFID 307 370
FT VARSPIC 386 387 SL -> NG (in isoform 2).
FT VARSPIC 388 466 /FTid=VSP_008631.
FT VARSPIC 344 372 Missing (in isoform 2).
FT VARSPIC 344 372 /FTid=VSP_008632.
FT VARSPIC 344 372 Missing (in isoform 3).
FT VARSPIC 344 372 /FTid=VSP_008633.
SQ SEQUENCE 466 AA; 51681 MW; B4312D1E885E386D CRC64;

Query Match 23.8%; Score 738; DB 1; Length 466;
Best Local Similarity 34.2%; Pred. No. 1.6e-33;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;

QY 34 RIVCYVGTWSVYHKVDP-YIYEDIDPKFCHLMYGPAKIDBYKYIQTVPFYQDDNHSW 92
DB 23 KLVCYFTNWAQYRQGEARFLPKDPLSLCHTLIYAFAGMTNQLST-----TEW 71
QY 93 -EKGYERFNNLRUKNPPELLTMI SLGWYEGSEKYSDMAANPTVYRQFIOSVLDFLQEYK 151
DB 72 NDELYQEENGLKKNPKLTKLLAIGWNFGTQKFTDMVATANNRQTFVNSAIRFLKYS 131
QY 152 FDGLDLWEYFGSRLGNPKIDKONYLALVRELKDAFEPHG-----YLLTAAVSPGOK 204
DB 132 FDGLDLWEYFGSQ-GSPAVDKERFTTLVQDLANAFQEAQTSCKERILLLSAAVPAQTY 190
QY 205 IDRAYDIKELNKLFDWNNVMTYDHGWENFYGHNAFLYKRPDTEDELHYFNNVYTMHY 264
DB 191 VDAGEYVDKIAQNLDVNLNAYDFHGWSEKVTGHSFLYKRQESGAAS-LNYDAAVQQ 249
QY 265 YLNNGATEDKLVMGVPFYGRAWSIEDRSKLLGDPAGKMSPPGISGEEGLVSLIELCOL 324
DB 250 WLQGTTPASKILGNPTVGRSFTLIASSSDTRVGAPATGSGTPGFTKEGGMLAYEVCWS 309
QY 325 FQKEEWHIQDYENYAPYGNDKITWGVYDLDIASISCKLAFELKELGELGVGMVWSLENDDFK 384

```

[illegible]


```

RX MEDLINE=98324849; PubMed=9662472;
RA de la Vega H., Specht C.A., Liu Y., Robbins P.W.;
RT "Chitinases are a multi-gene family in Aedes, Anopheles and
RL Drosophila."
RL Insect Mol. Biol. 7:233-239(1998).
CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -I- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -I- SIMILARITY: Contains 2 chitin-binding type-2 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; UA2895; AAA83586.1; -.
CC EMBL; AF026152; AAB81847.1; -.
CC DR PIR; T15408; T15408.
CC DR PIR; T37249; T37249.
CC DR WormPeP; C04F6.3; CE03923.
CC DR InterPro; IPR002557; Chitin_bind_PeRA.
CC DR InterPro; IPR01223; Glyco_hydro_18.
CC DR InterPro; IPR001579; Glyco_hydro_18AS.
CC DR Pfam; PF01607; CBM 14; 2.
CC DR Pfam; PF00704; Glyco_hydro_18; 1.
CC DR ProDom; PD000471; Glyco_hydro_18; 1.
CC DR SMART; SM00494; ChtBD2; 2.
CC DR SMART; SM00536; Glyco_18; 1.
CC DR PROSITE; PS05940; CHIT_BIND_II; 2.
CC DR PROSITE; PS01095; CHITINASE_18; 1.
CC DR Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Repeat.
KW FT DOMAIN 478 534 CHITIN-BINDING TYPE-2 1.
FT FT ACT SITE 179 179 PROTON DONOR (BY SIMILARITY).
FT FT DOMAIN 563 617 CHITIN-BINDING TYPE-2 1.
FT FT ACT SITE 435 478 THR-RICH.
FT FT DOMAIN 138 138 W -> C (IN REF. 2).
FT FT CONFLICT 163 165 IFF -> TTS (IN REF. 2).
FT FT CONFLICT 176 176 I -> L (IN REF. 2).
FT FT CONFLICT 176 176 I -> L (IN REF. 2).
SQ SEQUENCE 617 AA; 66857 MW; DDAID2AAAC0ES4DA CRC64;

Query Match 22.7%; Score 705.5; DB 1; Length 617;
Best Local Similarity 30.2%; Pred. No 1.3e-31;
Matches 178; Conservative 90; Mismatches 210; Indels 111; Gaps 19;

QY 37 CYVGTWSVYHK-VDPYTIEDIDPFKTHLMYGFAPKIDYKYTIQVDFPDQDHNHNSWEKR 95
Db 57 CYFTNWAQYRQGRAKVPEDYTGLCITHLFASGWNNA-DYTVRAYDP--ADLPNDWAGE 113
QY 96 G-YERNVLRLKNPELTWTLISGWTGEGSEKYSDMAANPTPYRQOFTQSVLDFLOEYKFDG 154
Db 114 GWYRRYVKLKVTDTQTLKLLSFGWSEFGTALFGMAASASRKRVFIDSATITFVRTWGF 173
QY 155 LLDLDWYPGSRGNPKIDKQNYLALVELKDAPEPHG-----YLLTAAVSPGKKIDR 207
Db 174 IDIDWYPSG-----ATDMANVVALVKELKAACESEAGSTGKDRLLVTAAVAAGPATIDA 228
QY 208 AYDIKEBLKFLDMWNVMTYDYHGGWENFYGHNAPLYKRPDETDELHFTYFNVNTMHHYLN 267
Db 229 GYDIPNLARNFDFILLMSYDFFCAMASLVGFNSPLVATTELPAAEWNGW-NVDSSARVWQ 287
QY 268 NGATROKLVNGVPFYGRAMSIDERSKLKLGDPKAGMSPPGFISEGGVLVSYELCQLFQK 327
Db 288 KGMPKEKIIIVGMPYTRGRTWLTNANASAINPFTSGSPAKITQYVQ-EAGVGAYFFCEMLN 346
QY 328 EEWHIQVDEYNNAPYGVNDKIWGYDDLASISCKLAFLKELGVSQVWVWSELENDDRKGHC 387
Db 347 GATR-VWDSQSQVPLYVQGNQWWSYDDEESFANKMAYVKREGYGGAFVWTLDFDDFNAGC 405
QY 388 GPKN-----PLLNKVNHNINGDEKNSFECIL---GPSTTTPT-----PTTTPTTP 429

```

```

Db      406 SNSNGQYPLISVIAKELGG-----VIIPKKGVTTAPTVAITVTGRPMTSAVTT 458
               ||::: : |::| |::|| |::|| |::|
Qy      430 TTTPTTSPPTTPPTPSPTT-
               ||:| |::| |::| |::| |::| |::|
Db      459 TTAATTTTTRAAITNTASNTNVCSKGSDGFVNNSNCGLFVLCLSSKSYSMSCPGLOYS 518
               ||:| |::| |::| |::| |::| |::| |::|
Qy      450 -----PTTTPSPTTPTTSTPTTPTTPTPAFTTGTSPPTTEHTSETPKVITYVDGH 503
               ||:| |::| |::| |::| |::| |::| |::|
Db      519 ASLXYCTTSAGSGSVITTRAPTTTTKSAPT--VTTTTRAPTTT-----TPAF----- 564
               ||:| |::| |::| |::| |::| |::| |::|
Qy      504 LIKCYKEGDIPHPTNIHKYLVCBFVGNGWVHIMPCCPGTIWCOKLTIC 552
               ||:| |::| |::| |::| |::| |::| |::|
Db      565 --KTCKDGFQVPSCDLKFIRC--VNG--ISYNFECPNGLSFHADTWMC 607
               ||:| |::| |::| |::| |::| |::| |::|

RESULT 7
OGP_PIG          STANDARD;              PRT;       527 AA.
ID_Q25990;
AC    01-NOV-1997 (Rel. 35, Created)
DT    01-NOV-1997 (Rel. 35, Last sequence update)
DI    28-FEB-2003 (Rel. 41, Last annotation update)
DE     Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE     (Oviductin) [Estrogen-dependent oviduct protein] (POSF-P3).
GN     OVGP1 OR OGP.
OS     Sus scrofa (Pig).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX     NCBI_TaxID=9823;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     STRAIN=Yorkshire X Duroc X Hampshire; TISSUE=Oviduct;
RX     MEDLINE=97107140; PubMed=894988;
RA     Buih W.C., Alvarez I.M., Choi I.; Cleaver B.D., Simmen F.A.;
RT     "Molecular cloning and characterization of an estrogen-dependent porcine oviductal secretory glycoprotein.";
RL     Biol. Reprod. 55:1305-1314(1996).
CC     -!- FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
CC     in the fertilization process and/or early embryonic development.
CC     -!- SUBCELLULAR LOCATION: Secretory granule.
CC     -!- TISSUE SPECIFICITY: Oviduct.
CC     -!- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
Cc      EMBL; U43490; AAA85445.1; -.
DR      InterPro; IPR001223; Glyco_hydro_18.
DR      InterPro; IPR001579; Glyco_hydro_18AS.
DR      Pfam; PF00704; Glyco_hydro_18; 1.
DR      ProDom; PD000471; Glyco_hydro_18; 1.
DR      SMART; SM00636; Glyco 18; 1.
DR      PROSITE; PS01095; CHITINASE 18; FALSE_NEG.
KW      Glycoprotein; Fertilization; Signal.
FT      SIGNAL                  1..21
FT      CHAIN                   22..527
FT      CARBOHYD                62..62
FT      CARBOHYD                402..402
FT      CARBOHYD                441..441
SQ      SEQUENCE 527 AA; 58519 MW; 31B78FA9CA2363A2 CRC64;
                                     BY SIMILARITY.
                                     OVIDUCT-SPECIFIC GLYCOPROTEIN.
                                     N-LINKED (GLCNAC...) (POTENTIAL).
                                     N-LINKED (GLCNAC...) (POTENTIAL).
                                     N-LINKED (GLCNAC...) (POTENTIAL).
Query Match           22.4%; Score 696.5; DB 1; Length 527;
Best Local Similarity 34.6%; Pred. No. 3.2e-31;
Matches 178; Conservative 86; Mismatches 192; Indels 59; Gaps 19;

Qy      10 IMACIGLNAGIKRDNDYSKNRMRIVCYTGWSYHKVDPYTI--BDIDFKCHLMYG 67
               ||:| |::| |::| |::| |::| |::| |::|

```

[illegible]

```

DR PIR; S57197; S57197.
DR InterPro: IPR001223; Glyco_hydro_18.
DR InterPro: IPR001579; Glyco_hydro_18AS.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR ProDom: PD000471; Glyco_hydro_18; 1.
DR SMART: SM00636; Glyco_18; 1.
DR PROSITE: PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 537
FT CARBOHYD 399 537
SQ SEQUENCE 537 AA; 59617 MW; CFCDEB6F0212D791 CRC64;

Query Match 22.1%; Score 685.5; DB 1; Length 537;
Best Local Similarity 33.3%; Pred. No. 1.3e-30;
Matches 173; Conservative 95; Mismatches 183; Indels 69; Gaps 19;

QY 10 IMACIGLMNASIKRDNDYSKNPMRVICVGVTSVYHKVDPTI--EDIDPKCTHLMYG 67
Db 1 LLLCVGLL---LVKHHGAAH--KLVCYFTNWA-FSRPGASILPRDLDPFLCTHLVFA 54
QY 68 FAKIDEYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYS 127
Db 55 FASMNQ--IVPKDP-QD-----EKILYPFNKLEKERNRGLKTLISGWNFTVFT 105
QY 128 DMAANPTYRQQTQSVLDLFLQYKFDGLDWEYFGSRLGNPKIDKQNYLALVRELKDAF 187
Db 106 TMLSTFNSRERFVSSVIALLRTHGFDGLDFFLYPGLR-GSPARDRWTFVFLLELLQAF 164
QY 188 EPHG-----VLLTAUSPGKDKIDRAVDIKELKLPDMVMWYDYHGGWENFYHNA 240
Db 165 KNEAQLTMRPLLGAASGDPHVVKAYEARLLGRLDFFISVLDLHGSWEKVTGHS 224
QY 241 PLVKKPDETDLHTFVNVTWHYLLNNGATDKLVMGVPFYGRAWSTEDRSKLGIDPA 300
Db 225 FLFSLPGDK-----SSAYNMYRWQLGVPPEKLMGLPTVGRTHLKLKASQNELRAQ 278
QY 301 KGMSPPGFTSGBEGLSYTELQCFQKEWHIQDYDEYNAPYGYNDKIWGVDDDLASIS 360
Db 279 VGPASPGKYTKAGFLAYVEICFVRRACKRWINDQY--VPYAFKGEWGVDDDAISFGY 336
QY 361 KLAFLKELGVGMVWVLENDLDFKGH-CGP-KNPLINKVHNMINDGNKSNFECILGPSTT 418
Db 337 KAFFIKREHFGAMVWTLDDLDFRGVFCGTGFPFLVHTLNLLVNDFFSS----- 386
QY 419 TTPP-----TTTPTPTTPTTPTTPTTPTTPTT-----TTTPTTPTTPTTPTT 466
Db 387 TSPKFWFTAVNSRIGEMPTMTDRDLTGLGILPGGEAVATHRKSETMTITPKGE 446
QY 467 TPTPT-TP-----TPAPTSTPSP-----TTTEHTSETP 494
Db 447 IATPTPTPLSFGHTAAPGKTESPGKEPLTTVGHVAVSP 486

RESULT 9
OGP SHEEP STANDARD; PRT; 539 AA.
AC Q28542; Q28543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Estrus-associated oviductal glycoprotein) (OEGP).
DE OVGP1 OR OGP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.

```

```

RC TISSUE=Oviduct;
RX MEDLINE=95269691; PubMed=7750470;
RA Desouza M.M., Murray M.X.;
RT "An estrogen-dependent secretory protein, which shares identity with
RT chitinases, is expressed in a temporally and regionally specific
RT manner in the sheep oviduct at the time of fertilization and embryo
RT development.";
RL Endocrinology 136:2485-2496 (1995).
RN [2]
RP SEQUENCE OF 10-539 FROM N.A.
RX STRAIN=Merino; TISSUE=Oviduct;
RX MEDLINE=96329120; PubMed=8726871;
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;
RT "Cloning and sequencing of a cDNA encoding an ovine
RT oestrus-associated oviductal protein.";
RL Reprod. Fertil. Dev. 8:305-310 (1996).
CC -!- FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
CC in the fertilization process and/or early embryonic development.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: Oviduct.
CC -!- DEVELOPMENTAL STAGE: Levels are highest in the fimbria and ampulla
CC at estrus and on day 1 of pregnancy, when gamete transport and
CC fertilization occurs in the E2-dominated fallopian tube. Levels
CC decline significantly on day 2 and undergo a further significant
CC reduction on day 3 of pregnancy coincident with transport of the
CC embryo from the oviduct to the uterus, a reproductive stage
CC associated with rising progesterone levels.
CC -!- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U16719; AAC48471.1; --
DR EMBL; U17988; AAB01052.1; --
DR PIR; I46470; I46470.
DR InterPro: IPR001223; Glyco_hydro_18.
DR InterPro: IPR001579; Glyco_hydro_18AS.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR ProDom: PD000471; Glyco_hydro_18; 1.
DR SMART: SM00636; Glyco_18; 1.
DR PROSITE: PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21
FT CHAIN 22 539
FT CARBOHYD 402 402
FT CONFLICT 13 13
FT CONFLICT 122 122
FT CONFLICT 282 282
FT CONFLICT 375 375
FT CONFLICT 484 484
FT CONFLICT 520 520
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;

Query Match 22.0%; Score 685; DB 1; Length 539;
Best Local Similarity 33.8%; Pred. No. 1.4e-30;
Matches 173; Conservative 90; Mismatches 185; Indels 64; Gaps 17;

QY 10 IMACIGLMNASIKRDNDYSKNPMRVICVGVTSVYHKVDPTI--EDIDPKCTHLMYG 67
Db 4 LLLVWVGLL---LMLKHHDGAAH--KLVCYFTNWA-FSRPGASILPRDLDPFLCTHLVFA 57
QY 68 FAKIDEYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYS 127
Db 58 FASMNQ--IVPKDP-LD-----EKILYPFNKLEKERNRGLKTLISGWNFTVFT 108
QY 128 DMAANPTYRQQTQSVLDLFLQYKFDGLDWEYFGSRLGNPKIDKQNYLALVRELKDAF 187
Db 109 KMLSTFNSRERFVSSVIALLRTHGFDGLDFFLYPGLR-GSPARDRWTFVFLLELLQAF 167

```



```
DR SMART, SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 721 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT DOMAIN 486 632 [TAP]-G-[IV].
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;

Query Match 21.9%; Score 680; DB 1; Length 721;
Best Local Similarity 33.7%; Pred. No. 3.6e-30;
Matches 175; Conservative 79; Mismatches 184; Indels 82; Gaps 16;

QY 34 RIVCVGTWVYHKVDYTI--EDIDPKCTHLMYGAIDYKYTIQVDPDPQDDNHS 91
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 23 KLVCYFTNWA-HSRPGPASPIMPHDLDFLCTHLIFAPASNNQI---VAKNLQDENVL- 77
QY 92 WEKGYERFNNLRKKNPELTMTISLGGWYEGSEKSYDMAANPTYRQOFLOSVDLQEQYK 151
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 78 -----YPEFNKLKERNRELKTLISIGGNWFGTGRFTAMLTLANREKFDVSLFRLHG 132
QY 152 FDGLDLQWYVPGSLGNPKIDKONYALVRELKDAPEPHG-----YLLTAAVSPGDK 204
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 133 FDGLDLFLFLPGLR-GSPHPRNWFLELBELOPAPFEREALLTQHPRLLLSAAVSGIPSI 191
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 IDRAYDIKELNKLFDWNVMTYDYGWENFYGNAPLYKRPDETDELHTYFNVNTMYH 264
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 IHTSYDALLGRRDLDFINLSYDLHGSWEXFTGHSPLFLSPEDSK-----SSAYAMNY 245
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 YLNNGATRDKLWGVPPYGRAMSTEDRSKULGPDAKMGSPPFISGEQVLSYIELCOL 324
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 246 WRKLGTPADKLIMFPYTYGRNFYLLKESKNGLOQTASMGSPASPGKYTKAQGLAYVEVCSF 305
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 FQKEWHIQDYENYAPGYNDKILWGYDDLASLKLFLKELGVGVWVWNLNDFFX 384
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 306 VQRAKKE--WIDYQVYPAFKGEWLGYDDTISFYKAMYKREHFGGANVWTLDMDDVR 363
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 385 G-HCG-PKNLNLKVNHNINGDEKNSPECILGPSTTTTP------422
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 364 GTFCNGGPFPLVHILNELLVQTESN-----TLPQFQFTSSVNASGPGSENTAL 413
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 423 ---TTTTPTPTTTTTPSPTPP-----TTTPS---PTTPTTPTPTTPTTPTPTPT 469
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 TEVLTITTIKILPGGEMATEVHRRYENMTVPSDGSVFGGTASPKHATVPENNTWA 473
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 PTTPTP-----APTTPSPPTT---EHTSETPKYTYV 500
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 AEAKTMTSLDFFSKTTTGVSKTITGSKTTTGVSKTITGV 513
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
OGP HUMAN STANDARD; PRT; 678 AA.
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Mucin 9).
GN OVGP1 OR OGP OR MUC9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=95119256; PubMed=7819450;
RA Arias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
```


AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein).
GN OVGPI OR OGP.
OS Papio anubis (olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Oviduct;
RC MEDLINE=98244335; PubMed=9584944;
RX Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
RA Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: characteristics of an oestradiol-dependent
RT oviduct-specific glycoprotein."
RL Hum. Reprod. Update 3:541-552(1997).
RN [2]
RN PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RP TISSUE=Oviduct;
RC MEDLINE=91367180; PubMed=1716345;
RX Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA Jaffe R.C.;
RT "cloning of a recombinant complementary DNA to a baboon (Papio
RT anubis) estradiol-dependent oviduct-specific glycoprotein."
RL Mol. Endocrinol. 5:356-364(1991).
CC -!- FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
CC in the fertilization process and/or early embryonic development.
CC -!- SUBCELLULAR LOCATION: Secretory granules.
CC -!- TISSUE SPECIFICITY: Oviduct.
CC -!- DEVELOPMENTAL STAGE: At the time of ovulation.
CC -!- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59903; AAB39765.1; -
DR PIR; A37954; A37954.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY
FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 623 AA; 69291 MW; 9E21CE481PFF1368 CRC64;

Query Match 21.1%; Score 654.5; DB 1; Length 623;
Best Local Similarity 30.7%; Pred. No. 7.6e-29;
Matches 181; Conservative 87; Mismatches 200; Indels 121; Gaps 20;

QY 10 IMACIGLMNASIKRDNDYSKNEPMRTVCYGVTSVYHKVDPTI--EDIDPPKCTHLMYG 67
DB 4 LLLVGLV---LVLKHDGAH--KLVCYTNWA-HSRPGPASILPHDLDPFLCTHLIFA 57

QY 68 FAKIDEYKTIQVDFPDYQDDNHNSEKRGYERENLELNKLPETMTISLGGWYEGSEKYS 127
DB 58 FASNNNNQI---YAKQIDQ-----EKILYPEFNKLERNELKTLISLGGWNNFGTSRT 108

QY 128 DMAANTPYRQOFTQSVLDFLQYKDFGLDLDWYPGSRLGNPKIDKONYALVRELKDAF 187

DB 109 TMLSTPANREXFIASVISLLRTHDFDGLDFFLPGLR-GSPMDRWTFFLELIELLFAF 167
QY 188 BPHG-----YLLTAAVSPGKIDRAYDIKELNKLFDMMNNVTYDYHGOWENFYGHNA 240
DB 168 RKEALLTWRPRLLSAAVSGVPHIVQTSYDVRFLGRLLDFINVLSDYDLHGSWEKFTGHS 227
QY 241 PLYKRPDETDELHTYFNNVNTMHTYLNNGATRDKLVMGVPPFYGRANSIEDRSKLGDPA 300
DB 228 PLSLPEDPK-----SSAYANNYWRKLGAPSEKLIMGIPTYGRTFRLLKASKNGLOATA 281
QY 301 KGMSPPGFISGEGVLSVIELCOLF--QKEBWHIQYDEYNNAPYGYNDKIWVGYYDDLASI 358
DB 282 IGPASPGKYTKOAGFLAYFEICSFVWGAKHW----IDYQVVPYANKGKWWGDDAISF 337
QY 359 SKLAFLKELGVSGVMVMSLENDDEKFG-HCGP-KNPLINKVHNNMINGDEKNS-----F 409
DB 338 SYKAWFIRREHFGGANVWLLDMDVDRGTFCGTGPPFLVYVWMDILVRAEFSSTSLPQFWL 397
QY 410 ECIIGPSTTP-----TPPTTPTTPT-----TPPTTPTTPT----- 430
DB 398 SSANSSSTDPERLAVTKAWTTDIKILPPGGEAGVTEHGKCNNTITPRVTIVTPTKET 457
QY 431 -----TPPTTPTTPT-----TPPTTPTTPT-----TPPTTPTTPT----- 465
DB 458 VSLGKHTVALGEKTEITGATMTSVGHQSWTPGKALTVPVGHQSELPGKKTLLTPVGHQSV 517
QY 466 TTPPT-----TPTAPTST-----PSPTTSTSTSTPKYTYVVDGH 503
DB 518 TTGQKTLISVGHVTPGKKTLLTPVGHQSVTPVSHQSVSPGGMWMTVPVH 566

Search completed: March 22, 2004, 06:53:04

Job time : 23.2868 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:39:53 ; Search time 111.679 Seconds
(without alignments)
1568.003 Million cell updates/sec

Title: US-09-662-293-15
Perfect score: 3107
Sequence: 1 MKTIYALLSIMACIGLMNAS.....IMPCPPQTWCQEKLTGICE 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTEMBL 25:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 3107 | 100.0 | 555 | 5 Q9U6R7 | Q9U6R7 dermatophag |
| 2 | 1008.5 | 32.5 | 525 | 5 Q44079 | Q44079 anopheles g |
| 3 | 942 | 30.3 | 431 | 5 Q81SH5 | Q81SH5 araneus ven |
| 4 | 869 | 28.0 | 1635 | 5 Q17412 | Q17412 aedes aegyp |
| 5 | 854.5 | 27.5 | 4498 | 5 Q9W223 | Q9W223 drosophila |
| 6 | 847 | 27.3 | 544 | 5 Q5GQC4 | Q5GQC4 bombyx mori |
| 7 | 842 | 27.1 | 2838 | 5 Q5MP05 | Q5MP05 tenebrio mo |
| 8 | 839 | 27.0 | 566 | 5 Q5WR52 | Q5WR52 bombyx mori |
| 9 | 837.5 | 27.0 | 467 | 5 Q15993 | Q15993 penaeus jap |
| 10 | 837.5 | 27.0 | 543 | 5 Q5GR93 | Q5GR93 bombyx mori |
| 11 | 837.5 | 27.0 | 543 | 5 Q5GV05 | Q5GV05 bombyx mori |
| 12 | 837.5 | 27.0 | 565 | 5 P50710 | P50710 bombyx mori |
| 13 | 837.5 | 27.0 | 595 | 5 Q5VFR3 | Q5VFR3 drosophila |
| 14 | 832 | 26.8 | 460 | 5 Q5W2M7 | Q5W2M7 drosophila |
| 15 | 831.5 | 26.8 | 467 | 5 Q5ITU3 | Q5ITU3 penaeus van |
| 16 | 829.5 | 26.7 | 565 | 5 Q5GP59 | Q5GP59 bombyx mand |

| | | | | | |
|----|-------|------|------|-----------|---------------------|
| 17 | 823.5 | 26.5 | 557 | 5 Q8WTK0 | Q8WTK0 chorisoneu |
| 18 | 823 | 26.5 | 488 | 13 Q9OW34 | Q9OW34 bufo japoni |
| 19 | 814.5 | 26.2 | 552 | 5 Q9GV44 | Q9GV44 spodoptera |
| 20 | 810 | 26.1 | 574 | 5 Q17411 | Q17411 aedes aegyp |
| 21 | 802 | 25.8 | 483 | 5 Q23737 | Q23737 chelonius sp |
| 22 | 798 | 25.7 | 572 | 5 Q26042 | Q26042 penaeus jap |
| 23 | 790 | 25.4 | 620 | 5 Q9Y0D4 | Q9Y0D4 penaeus mon |
| 24 | 787 | 25.3 | 553 | 5 P91731 | P91731 hyphantria |
| 25 | 780 | 25.1 | 470 | 13 Q803B7 | Q803B7 brachydanio |
| 26 | 762.5 | 24.5 | 474 | 5 Q86LZ2 | Q86LZ2 luzomyia 1 |
| 27 | 754 | 24.3 | 500 | 13 Q7ZV48 | Q7ZV48 brachydanio |
| 28 | 738 | 23.8 | 466 | 4 Q13231 | Q13231 homo sapien |
| 29 | 736.5 | 23.7 | 473 | 11 Q99PH2 | Q99PH2 mus musculu |
| 30 | 734.5 | 23.6 | 472 | 11 Q9JLN1 | Q9JLN1 mus musculu |
| 31 | 734.5 | 23.6 | 473 | 11 Q9D803 | Q9D803 mus musculu |
| 32 | 733 | 23.6 | 482 | 13 Q8AV87 | Q8AV87 gallus gall |
| 33 | 732.5 | 23.6 | 387 | 4 Q9H3V8 | Q9H3V8 homo sapien |
| 34 | 731.5 | 23.5 | 472 | 6 Q9SM17 | Q9SM17 bos taurus |
| 35 | 723.5 | 23.3 | 688 | 5 Q8MS85 | Q8MS85 drosophila |
| 36 | 722.5 | 23.3 | 1013 | 5 Q96OM0 | Q96OM0 drosophila |
| 37 | 717 | 23.1 | 527 | 5 P91773 | P91773 penaeus jap |
| 38 | 716 | 23.0 | 460 | 5 Q8WS95 | Q8WS95 glossina mo |
| 39 | 710 | 22.9 | 484 | 5 Q9W092 | Q9W092 drosophila |
| 40 | 708.5 | 22.8 | 476 | 4 Q9BZP6 | Q9BZP6 homo sapien |
| 41 | 703 | 22.6 | 929 | 5 Q8WV79 | Q8WV79 haemaphysal |
| 42 | 702.5 | 22.6 | 983 | 5 Q9VZV2 | Q9VZV2 drosophila |
| 43 | 691 | 22.2 | 462 | 5 Q9W2M6 | Q9W2M6 drosophila |
| 44 | 687 | 22.1 | 381 | 11 Q99J84 | Q99J84 mus musculu |
| 45 | 687 | 22.1 | 389 | 11 Q8BKL8 | Q8BKL8 mus musculu |

ALIGNMENTS

RESULT 1

Q9U6R7

ID Q9U6R7 PRELIMINARY; PRT; 555 AA.

AC Q9U6R7; (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE 98kDa HDM allergen.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]

RP SEQUENCE FROM N.A.

RA Weber E.R., Hunter S., Stedman K., McCall C.;

RT "Cloning and Characterization of a 98 kDa Allergen from

RL Dermatophagoides farinae.";

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF178772; AAD52672.1; -

DR GO; GO:0005576; C:extracellular; IEA.

DR GO; GO:0008061; F:chitin binding; IEA.

DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.

DR GO; GO:0006030; P:chitin metabolism; IEA.

DR InterPro; IPR002557; Chitin bind PefA.

DR InterPro; IPR001223; Glyco hydro 18.

DR InterPro; IPR001579; Glyco hydro 18AS.

DR Pfam; PF00704; Glyco hydro 18; 1_18AS.

DR ProDom; PD000471; Glyco hydro 18; 1.

DR SMART; SM00494; ChtBD2; 1.

DR SMART; SM00636; Glyco 18; 1.

DR PROSITE; PS01095; CHITINASE_18; 1.

KW Glycosidase; Hydrolase.

SQ SEQUENCE 555 AA; 63238 MW; 0E4564A1A459B30B CRC64;

Query Match 100.0%; Score 3107; DB 5; Length 555;

Best Local Similarity 100.0%; Pred. No. 2.9e-186;

Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIYAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGWSVYHKVDPTTIEDIDPFK 60
DB 1 MKTIYAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGWSVYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGAFAKIDYKTYIQVDFPYQDDNHNHNSWEKRGYERFNNLRKLNPELTMTISLGGWY 120
DB 61 CTHLMYGAFAKIDYKTYIQVDFPYQDDNHNHNSWEKRGYERFNNLRKLNPELTMTISLGGWY 120
QY 121 EGSEKYSMAANPTVROQFIQSVLDFLOEYKFDGLDWEYFSGRLGNPKDKONYLALV 180
DB 121 EGSEKYSMAANPTVROQFIQSVLDFLOEYKFDGLDWEYFSGRLGNPKDKONYLALV 180
QY 181 RELKDAFPHGYLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNA 240
DB 181 RELKDAFPHGYLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNA 240
QY 241 PLYKRPDTELTHTYFNNVNTMYYLNGATRDKLVMGVPPFYGRAWSIEDRSKLGDP 300
DB 241 PLYKRPDTELTHTYFNNVNTMYYLNGATRDKLVMGVPPFYGRAWSIEDRSKLGDP 300
QY 301 KMSPPGISBEGVLSYIELCOLFQKEWHIOYDEYNAPYGYNDKIWGYDGLASIS 360
DB 301 KMSPPGISBEGVLSYIELCOLFQKEWHIOYDEYNAPYGYNDKIWGYDGLASIS 360
QY 361 KLAPLKELGVSVMVWSLENDDEKHCPCPKPLNKNVNMINGDEKNSFECILGPSTTP 420
DB 361 KLAPLKELGVSVMVWSLENDDEKHCPCPKPLNKNVNMINGDEKNSFECILGPSTTP 420
QY 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
DB 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
QY 481 TSPPTTHTSTPKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEYFVNGGWWHIMPCP 540
DB 481 TSPPTTHTSTPKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEYFVNGGWWHIMPCP 540
QY 541 PGTINCOEKLTCIGE 555
DB 541 PGTINCOEKLTCIGE 555

RESULT 2

Q44079 ID O44079 PRELIMINARY; PRT; 525 AA.
AC O44079;
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase.
GN AGCHI-1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gut;
RX MEDLINE=98030563; PubMed=9360958;
RA Shen Z., Jacobs-Lorena M.;
RT "Characterization of a novel gut-specific chitinase gene from the
RT human malaria vector Anopheles gambiae.";
RL J. Biol. Chem. 272:28895-28900 (1997).
DR EMBL; AF008575; AAB87764.1; -.
DR PIR; T44445; T44445
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin_bind.Pera.
DR InterPro; IPR001223; Glyco_hydro.18.
DR InterPro; IPR001579; Glyco_hydro.18AS.

RESULT 3

Q8ISH5 ID Q8ISH5 PRELIMINARY; PRT; 431 AA.
AC Q8ISH5;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase.
OS Araneus ventricosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Araneidae; Araneus.
OX NCBI_TaxID=182803;
RN [1]
RP SEQUENCE FROM N.A.
RA Han J.H., Kim S.R., Sohn H.D., Jin B.R.;
RT "Molecular cloning of a cDNA encoding the chitinase from the spider,
RT Araneus ventricosus.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV120879; AAN39100.1; -.

DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro.18; 1.
DR ProDom; PD000471; Glyco_hydro.18; 1.
DR SMART; SM00494; ChtB2; 1.
DR SMART; SM00636; Glyco.18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 525 AA; 57211 MW; 3234360EFFF36165 CRC64;
Query Match 32.5%; Score 1008.5; DB 5; Length 525;
Best Local Similarity 38.4%; Pred. No. 4e-55;
Matches 211; Conservative 93; Mismatches 187; Indels 59; Gaps 13;
QY 8 LSIMACIGLMNASIKRD-HNDYSKNPMRIVCVGWSVYHKVD-PYTIEDIDPFKCTHLM 65
DB 5 VGVLVVAFAFAEPHKAASAEKKVVCVGVWVTRPGNGRYDIHIDFSLCTHLM 64
QY 66 YGFAKIDYKTYIQVDFPYQDDNHNHNSWEKRGYERFNNLRKLNPELTMTISLGGWYSGSK 125
DB 65 YGFFGINE-DATVRIIDPYLDLEEN-WGRGHIKRFVGLXNVGPGKTLAAGWNEGSK 122
QY 126 YSMAANPTVROQFIQSVLDFLOEYKFDGLDWEYFSGRLGNPKDKONYLALVRELKD 185
DB 123 PSMAASGELRKEFISDCVAFQORHGFDGIDLWEYPAQORDGNPLIDRDNHAQLVEEMRE 182
QY 186 AFPHGYLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYXR 245
DB 183 EFHYGLLLTAAYSVFSGAGSYDIPRISKGFHFLNNVMDYHGAWSYCGINAPLYRG 242
QY 246 PDTELTHTYFNNVNTMYYLNGATRDKLVMGVPPFYGRAWSIEDRSKLGDPKAGMSP 305
DB 243 SADTTRDLQINVASIHFWLAQCGCTGRKLVIGIPLYGRNFIASANTQIGAPTGGGT 302
QY 306 PGFISBEGVLSYIELCOLFQKEWHIOYDEYNAPYGYNDKIWGYDGLASISCKLAPL 365
DB 303 VGYTRPGVMGVNEFCEKLEATEANDLRWSEQQVYAVRNNQWGVYDGLRLSVQLKVKYL 362
QY 366 KELGVSGVMVWSLENDDEKHC-CPKNPLNKNVNMINGDEKNSFECILGPSTTPTPT 424
DB 363 LDQGLGAMVWSLETDGFLGCGGGRYPLMHSEIRSLVNGGT-----PSTTTMPPSV 413
QY 425 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 484
DB 414 APPT-----STVAPGTTTTTGTANFGTTQPPT--SDAPNHTTTSTTTEGNGPTTRPPSG 466
QY 485 TTTEHTSETPKYTYVVDGHLIKCY--KEGDIHPHTNIHKYLVCEYFVNGGWWH 535
DB 467 -----DG-----PCAGRYGFVHPHTNCARYICLTADTYVEFT----- 500
QY 536 IMPCPGTIW 545
DB 501 ---CPFGTLF 507

| | | |
|--|---|---------------|
| Q9W223 | PRELIMINARY; | PRT; 4498 AA. |
| Q9W223 | AC | |
| 01-MAY-2000 | (TREMELrel. 13, Created) | |
| 01-MAY-2000 | (TREMELrel. 13, Last sequence update) | |
| 01-OCT-2003 | (TREMELrel. 25, Last annotation update) | |
| CG2989 | protein. | |
| CG2989 | [1] | |
| Drosophila melanogaster (Fruit fly). | | |
| Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | |
| Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | |
| Ephydroidea; Drosophilidae; Drosophila. | | |
| NCBI_TaxID=7227; | | |
| SEQUENCE FROM N.A. | | |
| STRAIN=Berkelley; | | |
| MEDLINE=20196006; PubMed=10731132; | | |
| Adams M.D., Calniker S.E., Holt R.A., Evans C.A., Gocayne J.D., | | |
| Amanatides P.G., Scherer S.B., Li P.W., Hoskins R.A., Galie R.F., | | |
| George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | |
| Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | |
| Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., | | |
| Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., | | |
| Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., | | |
| Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., | | |
| Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S., | | |
| Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., | | |
| Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | |
| Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., | | |
| de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., | | |
| Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | |
| Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., | | |
| Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., | | |
| Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., | | |
| Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., | | |
| Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C., | | |
| Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., | | |
| Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., | | |
| Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., | | |
| Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., | | |
| Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., | | |
| Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L., | | |
| Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M., | | |
| Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., | | |
| Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H., | | |
| Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., | | |
| Spiet E., Spradling A.C., Stapleton M., Strong R., Sun E., | | |
| Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., | | |
| Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., | | |
| Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., | | |
| Ye J., Yeh R.-F., Zavert J.S., Zhan M., Zhang G., Zhao Q., Zheng L., | | |
| Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., | | |
| Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; | | |
| "The genome sequence of Drosophila melanogaster." | | |
| Science 287:2185-2195(2000). | | |
| EMBL; AE003449; AAP46544.1; - | | |
| FlyBase; FBgn0030171; CG2989. | | |
| GO; GO:0005576; C:extracellular; IEA. | | |
| GO; GO:0008061; F:chitin binding; IEA. | | |
| GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA. | | |
| GO; GO:0005975; P:carbohydrate metabolism; IEA. | | |
| GO; GO:0006030; P:pichitin metabolism; IEA. | | |
| InterPro; IPRO02557; Chitin bind Per. | | |
| InterPro; IPRO01223; Glyco Hydro_18. | | |
| InterPro; IPRO01579; Glyco_hydro_18AS. | | |
| Pfam; PF01607; CBM_14; 1. | | |
| Pfam; PF00704; Glyco hydro_18; 1. | | |
| ProDom; PD000471; Glyco hydro_18; 1. | | |
| SMART; SM00494; Cnbd2; 1. | | |
| SMART; SM00636; Glyco_18; 1. | | |
| PROSITE; PS01095; CHITINASE_18; 1. | | |
| Glycosidase; Hydrolase. | | |
| SEQUENCE 4498 AA; 493096 MW; ABE65CE0AAB25489 CRC64; | | |

```
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CEM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 544 AA; 61069 MW; 89872DFC1DA23753 CRC64;

Query Match 27.3%; Score 847; DB 5; Length 544;
Best Local Similarity 36.58; Pred. No. 5.2e-45;
Matches 207; Conservative 78; Mismatches 218; Indels 64; Gaps 19;

QY 1 MKTIYAILSLVA-CIGLMNASIKRHDNDYKSNMRIVCVYGVTSVYHK-VDPYTIIDIDP 58
DB 1 MRAIFATLAVLASCAALVQCADSR-----ARIVCFYNWAVRPGVGRGIEDIPV 51

QY 59 FKCHLMYGFADIDYKTYQVDPYODDNNHWEKRGYERFNNLRKLNPELTMTSLGG 118
DB 52 DLCTHLYISFIGVTERKSESVLLIDPELD-----VDKSGFRNFTSLRSKHPDVKFWAVGG 106

QY 119 WYEGSEKYSMAANPTVROFIQSVLDLFOEYKFDGLDLDWEYVPSR-LGNPKIDKQNYL 177
DB 107 WAEGGSKYSMAVQKSTRMSFIRSVDFLKYDFDGLDLDWEYVPGAADRGGSFSDKDEFL 166

QY 178 ALVRELKDAF--BPHGYLLTAAYSPGKIDRAYDIKELNKLFDMMNVMYDVGWENF 235
DB 167 YFQELKRAIRADRGWELTAAYPLANFRLMGEGYHPELCQELDAIHVMSYDURGNAGF 226

QY 236 YGHNAPLYKRPDELTHTYFNVNMYTHYLLNNGATRDKLVMGVPFYGRAWSIED----- 290
DB 227 ADVHSPLYKRPD-QWAEKLVNVDGLNLEEKGCPTNKLVGIPYGRSFTLSAGNNY 285

QY 291 -----RSKVLGDPKAGMSPPGFISGEGVLSYIELCOLFOKE--EWHIQVDEYNNAPY 342
DB 286 GLGTYYNKEAGGDPAPYTNATGF-----WAYEICTEVDADGSGWTWKWDFGKCPY 338

QY 343 GYNDKLVWGDDLASISCKLAFKELGVSGVMVWSLENDPDKHCKGPKPLNKNVNMN 402
DB 339 AYGTQWVGVEDRSVEIKNNWIKERKGYLGAMTWAIDMDDFKLGEEENPLIKLLH---- 394

QY 403 GDEKNSPECILGPSTT---TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 459
DB 395 ---KHMSTYTPPARTGHATPTPEW--ARPPSPSPSEGDPI-----PTTTTTTVKPTT 445

QY 460 TTPSPPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 518
DB 446 RTTARTP--TTTTVPHGTTEEDFDINRPEVEEPTENEVDNADV-CNSEDDYVPDKKE 502

QY 519 IHKYLVEFVNGGWVHIMPCPGTIW 545
DB 503 CSKYWRG--VNGEGVQ--FSCQPGTIF 525

RESULT 7
ID Q8MP05 PRELIMINARY; PRT; 2838 AA.
AC Q8MP05;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CHIT5.
OS Tenebrio molitor (Yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
RN [1]
RP SEQUENCE FROM N.A.
RA Royer V., Fraichard S., Bouhin H.;

RT "A Novel putative insect Chitinase with multiple catalytic domains :
hormonal regulation during metamorphosis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487081; CAD31740.4; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0008843; F:endochitinase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR InterPro; IPR000634; S/T dehydratase_BS.
DR Pfam; PF01607; CEM_14; 5.
DR ProDom; PD000471; Glyco_hydro_18; 5.
DR SMART; SM00494; ChtBD2; 5.
DR SMART; SM00636; Glyco_18; 5.
DR PROSITE; PS01095; CHITINASE_18; 3.
DR PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW Signal; Hydrolase; Glycosidase.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 2838 CHITINASE.
SQ SEQUENCE 2838 AA; 321407 MW; 608B3F2A8E98B9B7 CRC64;

Query Match 27.1%; Score 842; DB 5; Length 2838;
Best Local Similarity 32.2%; Pred. No. 7.8e-44;
Matches 198; Conservative 92; Mismatches 205; Indels 120; Gaps 21;

QY 31 NP---MRIVCVGVTSVYHK-VDPYTIIDIDPFKCHLMYGFADIDYKTYQVDPYQD 86
DB 1919 NPNSFKVVCVFTWAWYRQGGPKLPSDIDPDJCTHIVGVFAVLNGDQLIKKHDWTAD 1978

QY 87 DNHSWEKRGYERFNNLRKLNPELTMTSLGSGWEGS--BKYSDMAANPTVYQFIQSVDL 145
DB 1979 -----FDNKFEYKVTAFKAG--IKVLIAGWNDSAGDKYSRLVWNPSSRRRTAHVVD 2031

QY 146 FLOEYKFGDLDDWEYP-----GSLGNPKIDKQNYLALVRELKDAFPHGYLLTAAYSP 200
DB 2032 FIENNPFGLDLDWEYKPCWQVDCNKG--PASDKSPFAEFVKELHEAFTPKGMLLSAAVSP 2090

QY 201 GCKIDRAYDIKELNKLFDMMNVMYDVGWENFYGHNAPLYKRPDELTHTYFNVNY 260
DB 2091 SRVVVDAGYDPTSRYLDTAVMVCYDYGQWDXKITGHVAPWYAHPPDDVD---ATFTNF 2147

QY 261 TTHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPKAGMSPPGFISGEGVLSYIE 320
DB 2148 TIHWIEKGADEKKLVMGPMYQSPSLADNNQNGVNAPTYGGGEAGETARGFLSYIE 2207

QY 321 LCOLFOKEBWHIQVDEYNN-APYGVNDKIWGYDDLASISCKLAFKELGVSGVMVWSLE 379
DB 2208 ICTNVINKGTVWRDRRGRIQGYAHLRDQWVSFDDIGIRHKSEFIFRAMGLGGGIWALD 2267

QY 380 NDDFKGCHG-PKNPLLNKNVHNGDKNSPECILGPSTTPT---PTTPTT-TPTTTPT- 434
DB 2268 LDDFENLGCCEYPLLRITNRLVRYAKPDKCLGKASSKPTQKPTKPTKPTKPTKPTQ 2327

QY 435 --TPSPPTPTPTTPT-sptt-----PTTPTPTPT----- 459
DB 2328 KPTPKSPPTPPYEPKPKSPKPSYGTSTESPEFPMVMPDPSVFCGRLFVADEKKNQNYLNCQ 2387

QY 460 -----TTTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 486
DB 2388 GELQIQVCPNGLFWNRDCHDWFENTECHPDGTTTAAPTTTTQTLLEVEVEVEVEPVTPT 2447

QY 487 -----TEHTSETPKYTYTVVGHILKCYKEGDIPIHTNIHKYL-----VCEFVNGGW 533
DB 2448 AAPVGGTVDTEGAYKWCYFTNW--AWYRQGD-----GKYLQDIDASLCTHINYGFA 2498

QY 534 V---HIMCPCPGTIW 545
```

```
Db 2499 VLDGSTWTLKPHDSW 2513
RESULT 8
Q8WR52 PRELIMINARY; PRT; 566 AA.
AC Q8WR52;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CHIB4.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kinsu x Showa; TISSUE=integument;
RA Abdel-Banat B.M.A.; Koga D.;
RT "Alternative mRNA splicing generates heterogeneity within Bombyx mori
RL gene for chitinase.";
DR EMBL: AF455139; AAL51080.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0008843; F:endochitinase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Signal; Hydrolase; Glycosidase.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 566 CHITINASE.
SQ SEQUENCE 566 AA; 63465 MW; 43FC717B3F6917D0 CRC64;

Query Match 27.0%; Score 839; DB 5; Length 566;
Best Local Similarity 35.9%; Pred. No. 1.7e-44;
Matches 207; Conservative 78; Mismatches 209; Indels 82; Gaps 20;

Qy 1 MKTIYAILSIMA-CIGLMNASIKRDNDYSKNPRLVYGVGTSVYHK-VDPYTIEDIDP 58
Db 1 MEAIFATLAVLASCAALVQCADSR-----ARIVCFNSWAVYRPGVGRYGIEDIPV 51

Qy 59 FKCTHLMYGFAXIDYKTIQVDFPYQDNNHNSKEKGYERFNNLRKNPELTMTISLGG 118
Db 52 DLCTHLIYSFIVGTVEKTSSEVLIDPELD-----VDKSGFRNFTSLRSKHDPKVMVAVGG 106

Qy 119 WYEGSEKYSDMAANTYRQQTQSVDLQELQKPGDLQVDPGSR-LGNPKIDKONYL 177
Db 107 WAEGSKSHVVAQKSTMSFIRSVDFLKYDFGLDLQVDPGSR-LGNPKIDKONYL 166

Qy 178 ALVRELKDAF--EPHYGLTAAVSPGKDIDRAYDIKELNKLFDMMNMTYDYHGGWENF 235
Db 167 YFVQELKGAIFRAGRGWELTAAPLANFLRMGSHVPELQCELDIAHVMYSLDRGNWAGF 226

Qy 236 YGHNAPLYKRPDETLHTYFNNVTMMYLLNNGATRLKLVGVDFYGRANWIED----- 290
Db 227 ADVHSPLYKRPD-QWAVEKLNVDNGLMWEKGCPTNKLVVGFYGRSFTLSAGNNNY 285

Qy 291 -----RSKLKGLDPAKGNPPGFTSGEEGVLSYELCOLFQKE--EWHIQDYENAPY 342
Db 286 GLGTVINKEAGGDPAPYTNATGF-----WAYYEICTEVDADGSGWTKWDEFGKCPY 338
```

```
Qy 343 GYNDKIWGVYDDLASISCKLAPLKLGLGVGVWVNSLENDDFKGGHCPKPNLLKNVHMIN 402
Db 339 AYKGTQWGVYEDPRSVIEIKANWIKKGYLGAMTAWIDMDFKGLCGEENPLIKLHKMS 398

Qy 403 GDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 458
Db 399 -----TYTVPARTGHTTPTPEWARP-PTSPDPSGDPTPTTPTTPTTPTTPTTPTTPTT 444

Qy 459 PTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 510
Db 445 TTTTARPTT-----TTTKVPHGTTEEDFDINVRPEVELPT-ENEVDNADV-CNSE 493

Qy 511 GP-IPHTNIHKYLVCFFVNGVWVHIMPCTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 545
Db 494 DDVVPDKCKSKYWC--VNGEGVQ--PSCQPGTIF 525

RESULT 9
O15993 PRELIMINARY; PRT; 467 AA.
ID O15993;
AC O15993;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fjchi-3.
OS Penaeus japonicus (Kuruma prawn).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Marsupenaeus.
OX NCBI_TaxID=27405;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe T.; Kono M.; Aida K.; Nagasawa H.;
RT "Purification and molecular cloning of a chitinase expressed in the
RL hepatopancreas of the penaeid prawn Penaeus japonicus.";
DR EMBL: AB008027; BAA22854.1; -.
DR HSP; P07254; ICTN.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
SQ SEQUENCE 467 AA; 51765 MW; 499F7095774CA445 CRC64;

Query Match 27.0%; Score 837.5; DB 5; Length 467;
Best Local Similarity 34.8%; Pred. No. 1.7e-44;
Matches 183; Conservative 85; Mismatches 173; Indels 85; Gaps 14;

Qy 35 IVCVGVGTSVYHK-VDPYTIEDIDPEKCTHLMYGFAXIDYKTIQVDFPYQD--DNHNS 91
Db 1 MCVYFGSWAVYRGLGKFDVEDIDPKICTHIVGFGAGL-AADSSIRVDPWMLCNDYG- 58

Qy 92 WEKRGYERFNNLRKNPELTMTISLGGWYEGSEKYSDMAANTYRQQTQSVDLQELQVYK 151
Db 59 --KCAVDRFTALQQANLTKALLAVGWGNEGSPKYSKMAADPALNRKFTTSSIELLKKG 116

Qy 152 FDGLDLDWEYVPGRLGNPKIDKONYLALVRELKDAFEPHYGLTAAVSPGKDIDRAYDI 211
Db 117 FDGLDLDWEYVPGRLGNPKIDKONYLALVRELKDAFEPHYGLTAAVSPGKDIDRAYDI 175

Qy 212 KELNKLFDMMNMTYDYHGGWENFYGHNAPLYKRP-DETDELHTYFNNVTMMYLLNNGA 270
Db 176 FEISKSLDLINVMYSYDLHGAWDDTYTHQSGLYAHPLDEGDNL--YLNVDFAISYIEWKA 233
```



```

QY 271 TRDKLVGVPEYGRAMSIDESKLGDPAGKMSPPGFIISGEBGLSYIELCOLFOKEW 330
Db 234 RPOGIALGIPYGRWTLASQOETGYAPAPQGAAGDWTSPGMLGNYEICYNTQTDW 293
QY 331 HIQDEYVNAFYGV---NDKIVGWYDGLASISCKLAFKELGSGVMVMSLENDPKGHC 387
Db 294 TVDDDPAMHEFYAYVFPNNWICSYDAASVVTYAEYAKSGLAGTWVNSVEDDFRGLC 353
QY 388 GPKNPLANKVHNMINGDEKNSFECILGPGSTTPTPTPTPTPTPTPTPTPTPTPTPSP 447
Db 354 H-----NRKYLHI-----KIMVEVFGGSITP----- 376
QY 448 TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTP 507
Db 377 -----PPLPTTRDPSEPTTTTRAPP-----PG-----IHC 403
QY 508 YKEDIHPTNIHKLVCPEF-VNCGWVHIMPCCPTGIMCOEKLTC 552
Db 404 TTLGNLPPDLCTHYILCSLNTSGFDEKVEVCEPTLFPQSPYC 449

RESULT 10
Q9GR93 PRELIMINARY; PRT; 543 AA.
AC Q9GR93;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CH1B.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kinshu x Showa; TISSUE=Integument;
RA Abdel-Banat B.M.A., Koga D.;
RT "Molecular cloning of Bombyx mori chitinase cDNA: a unique insert of 9
RT base pairs reduced the apparent molecular mass of the encoded
RT protein."
RL Submitted (DSC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB052294; BAB2007.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0008843; F:endochitinase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco 18; 1.
DR PROSITE; PS01095; CHITINASE 18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 543 CHITINASE.
SQ SEQUENCE 543 AA; 60371 MW; C36C71E3B84DF88 CRC64;

Query Match
Best Local Similarity 35.9%; Pred. No. 2.1e-44; Length 543;
Matches 207; Conservative 78; Mismatches 208; Indels 83; Gaps 20;

QY 1 MKTIVAILISIMA-CIGLMNASIKRHDNDYKNPMRIVCYGTWGSVYHK-VDPYTIEDIP 58
Db 1 MRAIFATLAVLASCAALVQ-----SDSRARIVCYSNWAVYRPGVGRYGIEDIPV 50

```

```

QY 59 FKCTHLMVGFAXIDEVKYTIQVDFPDQDNHNSWEKRGVERFNNLRKLNPELTWTISLGG 118
Db 51 DICTHLIVSFIQVTEKSEVLIIIDPEL-----VDKSGFRNFTSRSRHPDVKFVWVGG 105
QY 119 WYEGSEKISDMAANFYRQOFTQSVLDLQEQYKFDGLDLDWEYPGSR-LGNFKIDKQNYL 177
Db 106 WAEGGSKTSHWVAQKSTRMSFIRSVVDFLKKYDFDGLDLDWEYPGAADGGFSFSDKDFL 165
QY 178 ALVRELKDAF--EPHGYLLTAASVPGKOKIDRAYDIKELNKLFDWNVMTYDYGHWENF 235
Db 166 YFVQELKRAFIRAGRWELTAAPLANFRLMBGYHVPCLQELDALHVMVSDLRGNWAGF 225
QY 236 YGHNAPLYKGPDETDELHTYFNVTMTYYLANGATRDKLVMGVFPYFGRAMSIED----- 290
Db 226 ADVHSFLYKRPD-QWAVEKLVNDGLNLWEKGCPTKLVVGIFFYGRSFTLSAGNNY 284
QY 291 -----RSKLLGDPAGKMSPPGFIISGEBGLSYIELCOLFOKE--EWHIOYDEVYNAFY 342
Db 285 GLGTINKRAGGDPAPYNTATG-----WAYEICEVDADGSGWTKKWDFFKCBY 337
QY 343 GYNDKIWGYDDLASISCKLAFKELGSGVMVMSLENDPKGHCQPKNPLANKVHNM 402
Db 338 AVKGTQWGYEDPRSVETKQWIKWIKGVLGANTWALDMDDFKGLCGEENPLKLLHRMS 397
QY 403 GDEKNSFECILGPGSTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 458
Db 398 -----THTVPPARTGHTTTPPEWARP-PSPTSDPSEGDPIPTTTTITVKPTT 443
QY 459 PTTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 510
Db 444 TRTTARPTT-----TTTKVPHGTTEEDFINRVEPELPT-ENEVDNADV-CNSE 492
QY 511 GP-IDPHTNIHKLVCPEFVNGWVHIMPCCPTGIMCOEKLTC 545
Db 493 DDYVDPKCEKSYKWC--VNCGVQ--FSCQPGTIF 524

RESULT 11
Q9GV05 PRELIMINARY; PRT; 543 AA.
AC Q9GV05;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Chitinase precursor (EC 3.2.1.14).
GN CHI.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kinshu x Showa hybrid; TISSUE=Integument;
RX WEDLINE-21124800; PubMed-11222960;
RA Abdel-Banat B.M.A., Koga D.;
RT "A genomic clone for a chitinase gene from the silkworm, Bombyx mori;
RT structural organization identifies functional motifs.";
RL Insect Biochem. Mol. Biol. 31:497-508 (2001).
DR EMBL; AB048355; BAB13481.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0008843; F:endochitinase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; CBM 14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco 18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 543 CHITINASE.
SQ SEQUENCE 543 AA; 60371 MW; C36C71E3B84DF88 CRC64;

```

```

DR SMART; SM00636; Glyco 18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 543 CHITINASE.
SQ SEQUENCE 543 AA; 60982 MW; DE29675D83AEC2EF CRC64;

Query Match 27.0%; Score 837.5; DB 5; Length 543;
Best Local Similarity 35.9%; Pred. No. 2.1e-44;
Matches 207; Conservative 78; Mismatches 208; Indels 83; Gaps 20;

Qy 1 MKTYYAILSMA-CIGLMNASIKKDHNDYSKNPRIVCYVGTWGSYHK-VDPYTIEDIP 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MRAIFATLVLASCAALVQ-----SDSRALIVCFSNWAVRPGVGRYGIEDIPV 50

Qy 59 FKCTHLMYGFAKIDBYKXITQVDPYQDDNHNSEKGYERFNLRKLNKPELTTMISLG 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 DLCTHLLIYSFIVGTGKSEVLIIDPELD-----VDKSGFRNFTSLRSKHPDKVFWAVGG 105

Qy 119 WVEGSEKYSDMAAPTYRQOFTQSVLDFLOBYKFDGLDWEYPGSR-LGNPKDKQNYL 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 106 WAEGSKYSNMVAOKSTWMSFIRSVDFLKKYDFDGLDWEYFGAADRGSGSFSDKKFL 165

Qy 178 ALVRELKDAF--EPHGILLTAASFGDKIDRAYDIKELNKLFDWMNVMTYDEYGGWENF 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 YFVQELKRAFIRACRGWELTAAPLANFRLMEGYHVPELQCELDAIHWMSYDLRGNWAGF 225

Qy 236 YGNAPLYKRDETDLHTYFNVNMYHYLLNNGATDKLVGVPPFYGRAWSIED----- 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 ADVHSPLYKRPHD-QWAYEKLNVNDGLNMLEKGCGPTNKLVGIPFYGRSFTLSAGNNY 284

Qy 291 -----RSKLGLDPAKMSPPGISGEGVLSVIELCOLPQKE--EWHIQDYFYKNAPY 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 GLGTYYINKEAGGGPPAPYTNATG-----WAYEICTEYDADGSGWTKWDFGKCPY 337

Qy 343 GYNDKIWGYDDLASICKLAFLELGVGVWWSLENDPFKGCGRKPNKLLNKHVMIN 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 AYKGTQWGYEDPSRVEIKMWIKEKGLGAMTWAIMDDPFKGLCGEENPLIKULHKHMS 397

Qy 403 GDEKNSFECILGPSITTTPTTTPTTTPTTTPTTTPTTTPTTTPTTTPTTTPTTTPTT 458
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 -----NVTVPARTGHTTTPPEWARP-PTSPSPSEGDPTPTTTTTVKTPTT 443

Qy 459 PTTTSPSTPTPTTPTAPTPTSTPTTTE-----HTSETPKYTYVDGHLIKCYKE 510
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 444 TRTTARPIT-----TTTKVPHGTTTEDFDINVRPEVEELPT-ENEVDNADV-CNSE 492

Qy 511 GD-IPHPTNIHKYLVCEPVGNGVWVHIMPCPGTIW 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 493 DDVYDPKKECKSYWRC--VNGEGVQ--FSCOPGTIF 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 12
P90710
AC AC PRELIMINARY; PRT; 565 AA.
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Chitinase-like protein.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
FX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98319808; PubMed=9654739;
RA Kim M.G., Shin S.W., Bas K.S., Kim S.C., Park H.Y.;
RT "Molecular cloning of chitinase cDNAs from the silkworm, Bombyx mori
and the fall webworm, Hyphantria cunea.";
RL Insect Biochem. Mol. Biol. 28:163-171(1998).
DR EMBl; U86876; AAB47538.1; -.

```

| | |
|----|--|
| DR | GO; GO:0005576; C:extracellular; IEA. |
| DR | GO; GO:0008061; F:chitin binding; IEA. |
| DR | GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA. |
| DR | GO; GO:0005975; P:carbohydrate metabolism; IEA. |
| DR | GO; GO:0006030; P:chitin metabolism; IEA. |
| DR | InterPro; IPR002557; Chitin_bind_ParA. |
| DR | InterPro; IPR001223; Glyco_Hydro_18. |
| DR | InterPro; IPR001579; Glyco_Hydro_18AS. |
| DR | Pfam; PF01607; CBM_14; 1. |
| DR | Pfam; PF00704; Glyco_hydro_18; 1. |
| DR | ProDom; PD000471; Glyco_hydro_18; 1. |
| DR | SMART; SM00494; ChtBD2_1. |
| DR | SMART; SM00636; Glyco_18; 1. |
| DR | PROSITE; PS01095; CHITINASE_18; 1. |
| KW | Glycosidase; Hydrolase. |
| SQ | SEQUENCE 565 AA; 63394 MW; 77D26D014875F3E2 CRC64; |

| | |
|---|---------------------------------------|
| Query Match | 27.08; Score 837.5; DB 5; Length 565; |
| Best Local Similarity | 35.94; Pred. No. 2.2e-44; |
| Matches 207; Conservative 78; Mismatches 208; Indels 83; Gaps 20; | |

| | | |
|----|--|-----|
| Qy | 1 MKTIYAILISMA-CIGLMMASIKRDHNDYSKNPMRIVCYVGTMSVHK-VDPYTIEDIP | 58 |
| Dd | 1 MRAIFATLAVLASCAALVQ-----SDSRARIVCFENWAVRPVGVRGIEDIPV | 50 |
| Qy | 59 FKCTHLMYGFAKIDYYKITTVQDFPVDNNHNWEKKGYERFNRLRKUNPELTMTISLG | 118 |
| Dd | 51 DLCTHLTYSFIGVTESSEVLIDPELD----VDKSGFRNFSLRSKHDPVFMVAVG | 105 |
| Qy | 119 WYGESEKYSDMAANPTVRQFIQSVDLFQEQKFEDGLDWEYEGSR-LGNPKIDQNKL | 177 |
| Dd | 106 WAEGGSYSMWQAOKSTRMFIRSVDFLKDYFDGLDWEYPGAADRGGSFSDDKFL | 165 |
| Qy | 178 ALVRELDAEF--EPHGVIALLTAAYSPGDKIDRAYDIKELNKLFDMMVMVTYDHGGWENF | 235 |
| Dd | 166 YFVOELKRATIRAGRGHELTAAVPLANFRLMEGHYVELCOELDAIHWNSYDLGNWAGP | 225 |
| Qy | 236 YGHNAPLYKRPDETDELHTYFNVNTHYYLLNNGATRDKLVMGVPFYCRAWSTED---- | 290 |
| Dd | 226 ADVHSPLYKRPD-QWAYEKLVNDGLNLWEKGCPNTKLWGVFFYGRSFTLSAGNNY | 284 |
| Qy | 291 -----RSKLGLDPAKGMPPGISGEBGVLYIELCOLFOKE--EWHIQYDEYYNAPY | 342 |
| Dd | 285 GLGIYINKAAGGDPAFYTNATGP-----WAYYEICTEVDADGSGWTKKWDEFKCPY | 337 |
| Qy | 343 GYNKDIXWGVDLDAISCKLAFLKELGVGVWWSLENDDFKHCGRPKNLLNVHNMN | 402 |
| Dd | 338 AYKTQWGVGEDPRSVAIKNNWIKKEYLGAANTWAIDMDDFKGLCGGENPLIKULHKMS | 397 |
| Qy | 403 GDEKNSPECILGSTTTTPTTTTPTTTTPTTTPPTSPTTPTT---TPSPTTTTPSPTT | 458 |
| Dd | 398 -----NYTVPPARTGHTTTPEWARP-PSTPSDPSSEGDPITTTTTTIVKPTT | 443 |
| Qy | 459 PTTPSPPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTT-----HTSETPKYTYVDGHLIKYKE | 510 |
| Dd | 444 TRTTARPTT-----TTTKVPHGTTTEEDFINVRPEVEELPT-EDEVNDADV-CNSE | 492 |
| Qy | 511 GD-I-PHPNTNIHKLVCEFFVNGGWVHWIMPCPGTIW | 545 |
| Dd | 493 DDYVFDKKECSKYWRCL-VNGEVGQ--FSCQPGTIF | 524 |

| | |
|--|---------------------------|
| RESULT 13 | |
| OSVF83 | |
| ID Q9VFR3 | PRELIMINARY; PRT; 595 AA. |
| AC Q9VFR3; | |
| DT 01-MAY-2000 (TrEMBLrel. 13, Created) | |
| DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update) | |
| DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update) | |
| DE CG9307 protein (IP08894P). | |
| GN CG9307. | |
| OS Drosophila melanogaster (Fruit fly). | |
| OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | |

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Perrier S., Fleischmann W.,
RA Posler C., Gabriella A.E., Garg N.S., Gelbart M.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jaimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulev G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AF003452; AAF46663.1; -.
DR HSSP; P07254; 1CTN.
DR FlyBase; FBgn0034580; CG9357.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; P:chitin binding; IEA.
DR GO; GO:0015798; P:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001179; FKBP_PPIase.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18S.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 460 AA; 51055 MW; 8EF4F126F1DD8D9B CRC64;

Query Match 26.8%; Score 832; DB 5; Length 460;
Best Local Similarity 33.7%; Pred. No. 3.7e-44;
Matches 183; Conservative 87; Mismatches 175; Indels 98; Gaps 11;

QY 17 MNASIKRDNHYSKPMRVCVVGWTSVYHK-VDPYTTIEDIDPFKCTHLMYGFPAKIDYK 75
DB 1 MAASSAQNS-SKN---VVCQGTWVSYPKGFGEIDIDPFKCTHLMYGFPAKIDYK 55

QY 76 YTIQVDFDYDDNHNHSEKRGYERFNRLKPKPELTMTISLGWYEGSEKYSMDAANPTY 135
DB 56 GQLRVIDAVLDLENS-GRGNTKSNFALKPKNPVKTIVAVGWNEGSKRSLVARDPSK 114

QY 136 RQOFIQSVLDLFQYKFGDGLDWEYPSRGLGNPKIDKQNYLALVRELKDAPEPHGYLLT 195
DB 115 REKFVDVDFRQRHGFQDGLDWEYPSRGLGNPKIDKQNYLALVRELKDAPEPHGYLLT 174

QY 196 AAVSPGOKIDRAYDIKELNKLFDWNVNTYDHGWNFYGHNPY---KRPDETDL 252
DB 175 AAVGAQFAEISYDIPANVPVLDLINVAYDLHGPQDQVGINAPLYAAEKDASDSSGR 234

QY 253 HYFENVNTWYHYLLNNGATDKLVMGVPYFGWANSIEDRSKLKLGDPKAGHSPPGFISGE 312

DB 235 QOQLNVDAVVKYWLKAGAPAEKILGVPPYGRSFTLATAGNQPGAPHIGKGIAGNYSRE 294
QY 313 EGVLSIIELCQIFQKEWHIQDEYNAPYGNKDIWGYDGLASICKLAFELKELGVSG 372
DB 295 PGVLGNELCEMERBEWTKWEATQOQVEYAYRQRWGVEDFRSLAKAQYVMDNLGG 354
QY 373 VMVNSLENDFKGHCPK-NPILNKVHNMINDEKNSFECILGPSTTTTPTTPTTPTT 431
DB 355 IMIWSLESDDFRGTCCQGPYLLHEINRVLFGG-----NTPSLGTTESNR 399
QY 432 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 491
DB 388 -----NTPSLGTTESNR 399

QY 492 ETPK--YTTVDGHLIKYKGBDIPHTNIHKYLVCFVNGWVHINCPPTGTTWCQSK 549
DB 400 ESESEGFSCPAD-----PAGVIRPDNCKSFYVC-----SGKTHNFCPSGLNFDLDT 449

QY 550 LTC 552
DB 450 KSC 452

RESULT 15
Q8ITU3 ID Q8ITU3 PRELIMINARY; PRT; 467 AA.
AC Q8ITU3; DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chitinase.
OS Penaeus vannamei (Penaeid shrimp) (European white shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Litopenaeus.
OC NCBI_TaxID=6689;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang R.C.;
RT Cloning and characterization of a cDNA encoding a chitinase from
RT hepatopancreas of the Penaeus vannamei (Crustacea, Decapoda).";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF315689; AAN74647.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; P:chitin binding; IEA.
DR GO; GO:0015787; P:hydrolase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18S.
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; 1.
SQ SEQUENCE 467 AA; 51958 MW; 734A830C6F47F4CD CRC64;

Query Match 26.8%; Score 831.5; DB 5; Length 467;
Best Local Similarity 34.2%; Pred. No. 4.1e-44;
Matches 180; Conservative 87; Mismatches 174; Indels 85; Gaps 14;

QY 35 IVCVVGWTSVYHK-VDPYTTIEDIDPFKCTHLMYGFPAKIDYKTYIOVFDPYQD--DNHNS 91
DB 1 MCVCFGSAAYRQGLGKFDVEDIDPKICTHIFGFAGL-AHDSIRVLDPWNLCDNYG- 58

QY 92 WEKRGIERFNRLKPKPELTMTISLGWYEGSEKYSMDAANPTYRQOFIQSVLDLFQYK 151
DB 59 --KCAVDRFTALKQONANKAILAVGWNEGSPKYSKMAADPVLNRNRTSSIELKKG 116

QY 152 FDGLDWEYPSRGLGNPKIDKQNYLALVRELKDAPEPHGYLLTAAVSPGKDKIDRAYDI 211

Search completed: March 22, 2004, 06:59:24
Job time : 113.679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:30:23 ; Search time 168.752 Seconds
(without alignments)
929.256 Million cell updates/sec

Title: US-09-662-293-15

Perfect score: 3107
Sequence: 1 MKTIYAILSIMAGIGLWNAS.....IMPCPPGTWCQKLTCTIGE 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Van04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 3107 | 100.0 | 555 | 3 AAY52523 | Aay52523 House dus |
| 2 | 3107 | 100.0 | 555 | 5 AAU96327 | Aau96327 Der HW-m |
| 3 | 3107 | 100.0 | 555 | 5 AAU96328 | Aau96328 Der HW-m |
| 4 | 3014 | 97.0 | 536 | 3 AAY52525 | Aay52525 House dus |
| 5 | 3014 | 97.0 | 536 | 5 AAU96329 | Aau96329 Der HW-m |
| 6 | 2542 | 81.8 | 509 | 3 AAY52533 | Aay52533 D. pteron |
| 7 | 2542 | 81.8 | 509 | 5 AAU96337 | Aau96337 Der HW-m |
| 8 | 2542 | 81.8 | 509 | 5 AAU96338 | Aau96338 Der HW-m |
| 9 | 2475 | 79.7 | 490 | 3 AAY52535 | Aay52535 D. pteron |
| 10 | 2475 | 79.7 | 490 | 5 AAU96339 | Aau96339 Der HW-m |
| 11 | 1008.5 | 32.5 | 525 | 6 ABP72636 | Abp72636 Anopheles |
| 12 | 854.5 | 27.5 | 498 | 4 ABB58595 | Abb58595 Drosophil |
| 13 | 838.5 | 27.0 | 554 | 2 AAU01824 | Aau01824 Manduca s |
| 14 | 838.5 | 27.0 | 554 | 3 AAB07183 | Aab07183 Manduca s |
| 15 | 838.5 | 27.0 | 554 | 6 ABP72619 | Abp72619 Manduca s |
| 16 | 837.5 | 27.0 | 467 | 6 ABP72634 | Abp72634 Prawn chi |
| 17 | 837.5 | 27.0 | 585 | 6 ABP72625 | Abp72625 Bombyx mo |
| 18 | 837.5 | 27.0 | 595 | 4 ABB71737 | Abb71737 Drosophil |
| 19 | 832 | 26.8 | 460 | 4 ABB64366 | Abb64366 Drosophil |
| 20 | 815.5 | 26.2 | 583 | 5 AAE28197 | Aae28197 Flea chit |
| 21 | 812.5 | 26.2 | 635 | 5 AAE28203 | Aae28203 Flea chit |
| 22 | 811.5 | 26.1 | 559 | 5 AAE28199 | Aae28199 Flea PCFC |
| 23 | 810 | 26.1 | 574 | 6 ABP72635 | Abp72635 Aedes aeg |
| 24 | 802 | 25.8 | 483 | 6 ABP72633 | Abp72633 Chelonus |
| 25 | 787 | 25.3 | 553 | 6 ABP72626 | Abp72626 Hyphantri |

| | | | | | |
|----|-------|------|-----|------------|--------------------|
| 26 | 738 | 23.8 | 466 | 2 AAW08584 | Aaw08584 Human 50 |
| 27 | 738 | 23.8 | 466 | 2 AAW40259 | Aaw40259 Human chi |
| 28 | 738 | 23.8 | 466 | 2 AAY42425 | Aay42425 MO-218 cl |
| 29 | 738 | 23.8 | 466 | 4 AAE00432 | Aae00432 Human chi |
| 30 | 738 | 23.8 | 466 | 5 AAE25903 | Aae25903 Human chi |
| 31 | 738 | 23.8 | 466 | 5 ABB76291 | Abb76291 Human chi |
| 32 | 737.5 | 23.7 | 520 | 6 ABU09914 | Abu09914 Partial m |
| 33 | 736.5 | 23.7 | 473 | 7 ABR55543 | Abr55543 Amino aci |
| 34 | 735.5 | 23.7 | 459 | 7 ADC24231 | Adc24231 Human NOV |
| 35 | 734.5 | 23.6 | 473 | 7 ADC51464 | Adc51464 Chitotria |
| 36 | 732.5 | 23.6 | 387 | 2 AAW08585 | Aaw08585 Human 39 |
| 37 | 732 | 23.6 | 447 | 7 ADC24237 | Adc24237 Human NOV |
| 38 | 732 | 23.6 | 466 | 2 AAW40260 | Aaw40260 Human chi |
| 39 | 732 | 23.6 | 466 | 2 AAY42426 | Aay42426 MO-13B cl |
| 40 | 732 | 23.6 | 466 | 4 AAE00433 | Aae00433 Human chi |
| 41 | 732 | 23.6 | 466 | 5 AAE25904 | Aae25904 Human chi |
| 42 | 732 | 23.6 | 466 | 5 ABB76292 | Abb76292 Human chi |
| 43 | 731 | 23.5 | 466 | 2 AAW31498 | Aaw31498 Human chi |
| 44 | 729.5 | 23.5 | 452 | 7 ABR55544 | Abr55544 Amino aci |
| 45 | 723 | 23.3 | 466 | 6 ABP72621 | Abp72621 Human chi |

ALIGNMENTS

RESULT 1
AAY52523
ID AAY52523 standard; protein; 555 AA.
XX
AC AAY52523;
XX
DT 22-FEB-2000 (first entry)
XX
DE House dust mite (D. farinae) mite allergen protein (map) PDerf98-555.
XX
KW Mite allergen protein; map; high molecular weight; HMW-map; allergy;
KW house dust mite; IGE, immunoglobulin E; allergen; mapA; mapB;
KW hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
KW canine; veterinary; antibody; vaccine; immunisation.
XX
OS Dermatophagoides farinae.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note= "Signal peptide"
FT Protein 20..555
FT Protein /note= "Mature PDerf98-555"
XX
FN WO9954349-A2.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008524.
XX
PR 17-APR-1998; 98US-00062013.
PR 13-MAY-1998; 98US-0085295P.
PR 02-SEP-1998; 98US-0098909P.
XX
PA (HESK-) HESKA CORP.
XX
PI Mccall CA, Hunter SW, Weber ER;
XX
DR WPI; 2000-052700/04.
DR N-PSDB; AAZ38575, AAZ38576, AAZ38577, AAZ38578.
XX
FT Novel high molecular weight Dermatophagoides nucleic acid polypeptides
FT used to modify an animals' hypersensitivity to mite allergens.
XX
PS Claim 3; Page 111-113; 154pp; English.
XX
CC This sequence represents Dermatophagoides farinae mite allergen protein
CC (map) PDerf98-555. PDerf98-555 has a molecular weight of 98 kD,
CC comprising 555 amino acids, and is a component of the Dermatophagoides

CC farinae high molecular weight mite allergen protein (HWM-map)
CC composition. The HWM-map composition was isolated from a D. farinae
CC homogenate by gel filtration, with each fraction being analysed for the
CC presence of proteins that bound to IgE present in mite-allergic dog
CC antisera. Mite allergenic proteins and peptides, and nucleic acids
CC encoding them, may be used in therapeutic compositions to modify an
CC animal's hypersensitivity reaction to mite allergens. Animals that may be
CC treated include mammals and birds, especially felines, canines, equines,
CC humans, other pets, and work or domestic animals. The proteins or
CC fragments may also be used to diagnose allergies via a skin test. The
CC proteins and peptides can also be used to raise antibodies, which have a
CC variety of potential uses. For example, they can be used as vaccines to
CC passively immunise animals against dust mite hypersensitivity, as
CC positive controls in test kits and as tools to recover desired dust mite
CC allergens from a mixture of proteins
XX
SQ Sequence 555 AA;

Query Match 100.0%; Score 3107; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.8e-217;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKTIYAILSIMACIGLMNASIKRDHNDYSKNPMRIVCVGTWVSYYHKVDPTTIEDIDPFK 60
Db 1 MKTIYAILSIMACIGLMNASIKRDHNDYSKNPMRIVCVGTWVSYYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGFPAKIDYKYTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120
Db 61 CTHLMYGFPAKIDYKYTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120
QY 121 EGSKYSDMAANPYRQOFTOSVLDFTLQYKFDGLDWEYPSGRGNPKIDKQNYLALV 180
Db 121 EGSKYSDMAANPYRQOFTOSVLDFTLQYKFDGLDWEYPSGRGNPKIDKQNYLALV 180
QY 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 240
Db 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 240
QY 241 PLYKRPDETDLHTYFNNVTMYYLNGATRDKLVMGVFFYGRAWSIEDRSKLGDDPA 300
Db 241 PLYKRPDETDLHTYFNNVTMYYLNGATRDKLVMGVFFYGRAWSIEDRSKLGDDPA 300
QY 301 KGMSPPGFISGEEGVLSELCQFQKEWHIQYDEYNNAPYGYNDKIWGYDDLASISC 360
Db 301 KGMSPPGFISGEEGVLSELCQFQKEWHIQYDEYNNAPYGYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSGVMWSLENDDFKGGCPKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 420
Db 361 KLAFLKELGVSGVMWSLENDDFKGGCPKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 420
QY 421 TPPTTPT 480
Db 421 TPPTTPT 480
QY 481 TPSPTTTEHTSETPKYTYVDGHLIKCVKEGDIHPHTNIHKLVCFVNGGWVHMPCP 540
Db 481 TPSPTTTEHTSETPKYTYVDGHLIKCVKEGDIHPHTNIHKLVCFVNGGWVHMPCP 540
QY 541 PGTIWCQEKLTICIG 555
Db 541 PGTIWCQEKLTICIG 555

RESULT 2

AAU96327
ID AAU96327 standard; protein; 555 AA.

XX
AC AAU96327;

XX
DT 15-JUL-2002 (first entry)

XX
DE Der HWM-map polypeptide #14.

KW Der HWM-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
OS Dermatophagoides farinae.
XX
PN WO200222807-A2.
XX
XX 21-MAR-2002.
XX
PD 14-SEP-2001; 2001WO-US028730.
XX
XX 14-SEP-2000; 2000US-00662293.
XX
XX (HESK-) HESKA CORP.
XX
XX McCall CA, Hunter SW, Weber ER;
XX
XX WPI; 2002-351888/38.
XX
XX N-PSDB; ABK69571.
XX
PT New mite allergenic protein isolated from Dermatophagoides, designated
PT Der HWM-map protein, useful as a vaccine for treating mite allergy.
PS Claim 12; Page 114-116; 16:pp; English.
XX
CC The invention relates to an isolated mite allergenic protein of
CC Dermatophagoides, designated Der HWM-map protein, and its related nucleic
CC acid. The Der HWM-map protein is useful for eliciting an immune response
CC against Der HWM-map protein. The protein or a reagent comprising a non-
CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
CC cat) susceptible to or having an allergic response to a mite. A
CC therapeutic composition is useful for desensitising a host animal to an
CC allergic response to a mite. The DNA and protein can be used in the
CC detection of anti-Der HWM-map antibodies in animal fluids, and inhibition
CC of immunoglobulin (Ig)E or Der HWM-map protein activity associated with a
CC disease. Antibodies that bind to Der HWM-map are useful for inhibiting
CC binding of proteins to IgE, to prevent immunocomplex formation, thus
CC reducing hypersensitivity responses to mite allergens, and as vaccines
CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
CC represent Der HWM-map polypeptides of the invention
XX
SQ Sequence 555 AA;
Query Match 100.0%; Score 3107; DB 5; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.8e-217;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKTIYAILSIMACIGLMNASIKRDHNDYSKNPMRIVCVGTWVSYYHKVDPTTIEDIDPFK 60
Db 1 MKTIYAILSIMACIGLMNASIKRDHNDYSKNPMRIVCVGTWVSYYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGFPAKIDYKYTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120
Db 61 CTHLMYGFPAKIDYKYTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120
QY 121 EGSKYSDMAANPYRQOFTOSVLDFTLQYKFDGLDWEYPSGRGNPKIDKQNYLALV 180
Db 121 EGSKYSDMAANPYRQOFTOSVLDFTLQYKFDGLDWEYPSGRGNPKIDKQNYLALV 180
QY 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 240
Db 181 RELKDAPEPHGYLLTAAVSPGKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 240
QY 241 PLYKRPDETDLHTYFNNVTMYYLNGATRDKLVMGVFFYGRAWSIEDRSKLGDDPA 300
Db 241 PLYKRPDETDLHTYFNNVTMYYLNGATRDKLVMGVFFYGRAWSIEDRSKLGDDPA 300
QY 301 KGMSPPGFISGEEGVLSELCQFQKEWHIQYDEYNNAPYGYNDKIWGYDDLASISC 360
Db 301 KGMSPPGFISGEEGVLSELCQFQKEWHIQYDEYNNAPYGYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSGVMWSLENDDFKGGCPKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 420
Db 361 KLAFLKELGVSGVMWSLENDDFKGGCPKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNA 420

DR WPI; 2000-052700/04.
DR N-PSDB; AAZ38579, AAZ38580.
XX Novel high molecular weight Dermatophagoides nucleic acid polypeptides
PT used to modify an animals' hypersensitivity to mite allergens.
XX
XX Claim 3; Page 125-127; 154pp; English.
XX
XX This sequence represents Dermatophagoides farinae mite allergen protein
CC (map) pDerf98-536, the mature form of pDerf98-555 (AAV52523). pDerf98-536
CC has a molecular weight of 98 kD, comprising 536 amino acids, and is a
CC component of the Dermatophagoides farinae high molecular weight mite
CC allergen protein (HWM-map) composition. The HWM-map composition was
CC isolated from a D. farinae homogenate by gel filtration, with each
CC fraction being analysed for the presence of proteins that bound to IgE
CC present in mite-allergic dog antisera. Mite allergenic proteins and
CC peptides, and nucleic acids encoding them, may be used in therapeutic
CC compositions to modify an animal's hypersensitivity reaction to mite
CC allergens. Animals that may be treated include mammals and birds,
CC especially felines, canines, equines, humans, other pets, and work or
CC domestic animals. The proteins or fragments may also be used to diagnose
CC allergies via a skin test. The proteins and peptides can also be used to
CC raise antibodies, which have a variety of potential uses. For example,
CC they can be used as vaccines to passively immunise animals against dust
CC mite hypersensitivity, as positive controls in test kits and as tools to
CC recover desired dust mite allergens from a mixture of proteins
XX
XX Sequence 536 AA;
XX
XX Query Match 97.0%; Score 3014; DB 3; Length 536;
XX Best Local Similarity 100.0%; Pred. No. 1e-210;
XX Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 SIKRDHNDYKSNPRIVCYVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKYTIQ 79
Db 1 SIKRDHNDYKSNPRIVCYVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKYTIQ 60
QY 80 VFDPYQDDNHNWSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMDAANPTYRQOF 139
Db 61 VFDPYQDDNHNWSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMDAANPTYRQOF 120
QY 140 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTTAAVS 199
Db 121 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTTAAVS 180
QY 200 PGKDKIDRAYDIKELNKLFDWMVNTYDYGWENFYGHNAPLYKRPDETDELHTYFNVN 259
Db 181 PGKDKIDRAYDIKELNKLFDWMVNTYDYGWENFYGHNAPLYKRPDETDELHTYFNVN 240
QY 260 YTHMYLLNNGATRDKLVNGVVPFYGRAWSIEDRSKLGDPAGKMSPPGFISSGEEGVLSYI 319
Db 241 YTHMYLLNNGATRDKLVNGVVPFYGRAWSIEDRSKLGDPAGKMSPPGFISSGEEGVLSYI 300
QY 320 ELCQLFQKEEWHIOYDEVYNAPYGVNDKIWGYDDLASISCKLAFLKELGVSQVWMSLE 379
Db 301 ELCQLFQKEEWHIOYDEVYNAPYGVNDKIWGYDDLASISCKLAFLKELGVSQVWMSLE 360
QY 380 NDDFKHGCGPKNPLLNKVNHNMDKNSFECILGPSTTPTPTTPTTPTTPTTPTTPTTPT 439
Db 361 NDDFKHGCGPKNPLLNKVNHNMDKNSFECILGPSTTPTPTTPTTPTTPTTPTTPTTPT 420
QY 440 TPTTTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
Db 421 TPTTTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
QY 500 VDGHILKCYKEGDIHPHTNIHKYLCVFVNGWVWHIMPCPGTINQOEKLTICGE 555
Db 481 VDGHILKCYKEGDIHPHTNIHKYLCVFVNGWVWHIMPCPGTINQOEKLTICGE 536
RESULT 5
AAU96329
ID AAU96329 standard; protein; 536 AA.

XX AAU96329;
XX AC
XX 15-JUL-2002 (first entry)
XX DT
XX Der HWM-map polypeptide #16.
XX DE
XX
XX Der HWM-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
XX Dermatophagoides farinae.
OS
WO200222807-A2.
PN
XX 21-MAR-2002.
XX
XX 14-SEP-2001; 2001WO-US028730.
XX
XX 14-SEP-2000; 2000US-00662293.
XX
XX (HESK-) HESKA CORP.
XX
XX McCall CA, Hunter SW, Weber ER;
PI
XX WPI: 2002-351888/38.
XX
XX N-PSDB; ABK69575.
XX
XX New mite allergenic protein isolated from Dermatophagoides, designated
PT Der HWM-map protein, useful as a vaccine for treating mite allergy.
XX
XX Claim 12; Page 125-127; 161pp; English.
XX
XX The invention relates to an isolated mite allergenic protein of
CC Dermatophagoides, designated Der HWM-map protein, and its related nucleic
CC acid. The Der HWM-map protein is useful for eliciting an immune response
CC against Der HWM-map protein. The protein or a reagent comprising a non-
CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
CC cat) susceptible to or having an allergic response to a mite. A
CC therapeutic composition is useful for desensitising a host animal to an
CC allergic response to a mite. The DNA and protein can be used in the
CC detection of anti-Der HWM-map antibodies in animal fluids, and inhibition
CC of immunoglobulin (Ig)E or Der HWM-map protein activity associated with a
CC disease. Antibodies that bind to Der HWM-map are useful for inhibiting
CC binding of proteins to IgE, to prevent immunocomplex formation, thus
CC reducing hypersensitivity responses to mite allergens, and as vaccines
CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
CC represent Der HWM-map polypeptides of the invention
XX
XX Sequence 536 AA;
XX
XX Query Match 97.0%; Score 3014; DB 5; Length 536;
XX Best Local Similarity 100.0%; Pred. No. 1e-210;
XX Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 SIKRDHNDYKSNPRIVCYVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKYTIQ 79
Db 1 SIKRDHNDYKSNPRIVCYVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKYTIQ 60
QY 80 VFDPYQDDNHNWSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMDAANPTYRQOF 139
Db 61 VFDPYQDDNHNWSEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMDAANPTYRQOF 120
QY 140 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTTAAVS 199
Db 121 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTTAAVS 180
QY 200 PGKDKIDRAYDIKELNKLFDWMVNTYDYGWENFYGHNAPLYKRPDETDELHTYFNVN 259
Db 181 PGKDKIDRAYDIKELNKLFDWMVNTYDYGWENFYGHNAPLYKRPDETDELHTYFNVN 240
QY 260 YTHMYLLNNGATRDKLVNGVVPFYGRAWSIEDRSKLGDPAGKMSPPGFISSGEEGVLSYI 319
Db
XX

Db 241 YTMHYIANNATKDLVMGVFFYGRAWSIEDRSKXKLGDPAGKMSPPGISGEGVLSYI 300
Qy 320 ELQQLFQKEBWHIOYDEYNAPYGYNDKIWGVYDDIASISCKLAFKELGSGVGMVWSLE 379
Db 301 ELQQLFQKEBWHIOYDEYNAPYGYNDKIWGVYDDIASISCKLAFKELGSGVGMVWSLE 360
Qy 380 NDDFKHCGKPNLLKVNHWINGDEKNSPECILGSPSTTPTTPTTPTTPTTPTTPTTPT 439
Db 361 NDDFKHCGKPNLLKVNHWINGDEKNSPECILGSPSTTPTTPTTPTTPTTPTTPTTPT 420
Qy 440 TPTTPT 499
Db 421 TPTTPT 480
Qy 500 VDGHILKYKEGDIPIHPTNIHKYLVCFFVNGVGMVHIMPCTTTCQEKLTICGE 555
Db 481 VDGHILKYKEGDIPIHPTNIHKYLVCFFVNGVGMVHIMPCTTTCQEKLTICGE 536

RESULT 6
AAV52533
ID AAV52533 standard; protein; 509 AA.
XX AAV52533;
AC AAV52533;
XX
DT 06-AUG-2003 (revised)
DT 22-FEB-2000 (first entry)
XX
DE D. pteronyssius 98 kD mite allergen protein (map) Pderp98-509.
XX
KW Mite allergen protein; map; high molecular weight; HMW-map; allergy;
KW house dust mite; IgE; immunoglobulin E; allergen; mapB;
KW hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
KW canine; veterinary; antibody; vaccine; immunisation.
XX
OS Dermatophagoides pteronyssinus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note= "Signal peptide"
FT Protein 20..509
FT Protein /note= "Mature Pderp98-509"
XX
PN W09954349-A2.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008524.
XX
PR 17-APR-1998; 98US-00062013.
PR 13-MAY-1998; 98US-00852952.
PR 02-SEP-1998; 98US-0098909P.
XX
PA (HESK-) HESKA CORP.
XX
PI McCall CA, Hunter SW, Weber ER;
XX
XX WPI; 2000-052700/04.
XX N-PSDB; AAZ38585, AAZ38586, AAZ38587, AAZ38588.
XX
PT Novel high molecular weight Dermatophagoides nucleic acid polypeptides
PT used to modify an animals' hypersensitivity to mite allergens.
XX
PS Claim 3; Page 134-136; 154pp; English.
XX
XX This sequence represents Dermatophagoides pteronyssius mite allergen
CC protein (map) Pderp98-509. Pderp98-509 has a molecular weight of 98 kD,
CC comprising 509 amino acids, and has a high degree of homology with the D.
CC farinae 98 kD allergen, mapB (AAZ52523). Nucleic acid molecules encoding
CC Pderp98-509 were isolated from a D. pteronyssius cDNA library by
CC hybridisation with a probe encoding the D. farinae high molecular weight
CC map (HMW-map) composition. Mite allergenic proteins and peptides, and
CC nucleic acids encoding them, may be used in therapeutic compositions to

CC modify an animal's hypersensitivity reaction to mite allergens. Animals
CC that may be treated include mammals and birds, especially felines, the
CC canines, equines, humans, other pets, and work or domestic animals. The
CC proteins or fragments may also be used to diagnose allergies via a skin
CC test. The proteins and peptides can also be used to raise antibodies,
CC which have a variety of potential uses. For example, they can be used as
CC vaccines to passively immunise animals against dust mite
CC hypersensitivity, as positive controls in test kits and as tools to
CC recover desired dust mite allergens from a mixture of proteins. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 509 AA;
Query Match 81.8%; Score 2542; DB 3; Length 509;
Best Local Similarity 81.7%; Pred. No. 2e-176;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;
Qy 1 MKTIYAILSIACIGLMAASIKRDHNDYKQPMRIVCYGTWVYHKVDPTTIEDIDPFK 60
Db 1 MKTTFALFCIWACIGLMAAATKRDHNNYSKNPMRIVCYGTWVYHKVDPTTIEDIDPFK 60
Qy 61 CTHLMYGFADKIDYKTIQVDFDYQDDNHNSEKRGYERFNNLRKLNKPELTMTSLGWMY 120
Db 61 CTHLMYGFADKIDYKTIQVDFDYQDDNHNSEKRGYERFNNLRKLNKPELTMTSLGWMY 120
Qy 121 EGSEKYSDMAANPTVYRQFIQSVLDFLOEYKFDGLDLDWEYFGSRLGNPKIDKQNYLALV 180
Db 121 EGSEKYSDMAANPTVYRQFVQSVLDFLOEYKFDGLDLDWEYFGSRLGNPKIDKQNYLTLV 180
Qy 181 RELKDAFPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDWNNVMTYDHHGWNENFYGHNA 240
Db 181 RELKEAFEPFGYLLTAAVSPGKDKIDVAYELKELNQLFDWNNVMTYDHHGWNENFYGHNA 240
Qy 241 PLYKRPDETDLHTYFNVTMYLNGATRDKLVMGVPPFYGRAWSIEDRSKVLGDPA 300
Db 241 PLYKRPDETDLHTYFNVTMYLNGATRDKLVMGVPPFYGRAWSIEDRSKVLGDPA 300
Qy 301 KGMSPPGFIISGEGVLSYIELCOLFQKEBWHIOYDEYNAPYGYNDKIWGVYDDIASISC 360
Db 301 KGMSPPGFIISGEGVLSYIELCOLFQKEBWHIOYDEYNAPYGYNDKIWGVYDDIASISC 360
Qy 361 KLAFKLKELGSGVGMVWSLENDDFKGHCGPKYKPNLKNVHNMINGDEKNSPECILGSPSTTPT 420
Db 361 KLAFKLKELGSGVGMVWSLENDDFKGHCGPKYKPNLKNVHNMINGDEKNSPECILGSPSTTPT 420
Qy 421 TPTTPT 480
Db 421 TPTTPT 434
Qy 481 TPTTPT 537
Db 435 ---PTTDTSTSETPKYTTIDGHILKYKQGYLPHTDVKHLYLVCVEYIATFNGGWWHIM 491
Qy 538 PCPRTIWCQEKLTICGE 555
Db 492 DCPKGRWHAATLKNCIQE 509
RESULT 7
AAU96337
ID AAU96337 standard; protein; 509 AA.
XX AAU96337;
AC AAU96337;
XX
DT 15-JUL-2002 (first entry)
XX
DE Der HMW-map polypeptide #24.
XX
KW Der HMW-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
OS Dermatophagoides farinae.


```
AAU96339
ID AAU96339 standard; protein; 490 AA.
XX
AC AAU96339;
XX
DT 15-JUL-2002 (first entry)
XX
DE Der HMW-map polypeptide #26.
XX
KW Der HMW-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
OS Dermatophagoides farinae.
XX
PN WO200222807-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028730.
XX
PR 14-SEP-2000; 2000US-00662293.
XX
PA (HESK-) HESKA CORP.
XX
PI McCall CA, Hunter SW, Weber ER;
XX
DR WPI; 2002-351888/38.
DR N-PSDB; ABK69585.
XX
PT New mite allergenic protein isolated from Dermatophagoides, designated
PT Der HMW-map protein, useful as a vaccine for treating mite allergy.
PS Claim 12; Page 144-146; 161pp; English.
XX
CC The invention relates to an isolated mite allergenic protein of
CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
CC acid. The Der HMW-map protein is useful for eliciting an immune response
CC against Der HMW-map protein. The protein or a reagent comprising a non-
CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
CC cat) susceptible to or having an allergic response to a mite. A
CC therapeutic composition is useful for desensitising a host animal to an
CC allergic response to a mite. The DNA and protein can be used in the
CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
CC binding of proteins to IgE, to prevent immunocomplex formation, thus
CC reducing hypersensitivity responses to mite allergens, and as vaccines
CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
CC represent Der HMW-map polypeptides of the invention
XX
SQ Sequence 490 AA;

Query Match 79.7%; Score 2475; DB 5; Length 490;
Best Local Similarity 82.5%; Pred. No. 1.5e-171;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;

QY 22 KRDHNDYKSNRIVCYVGTWSVHKVDPTIEDIDPKCTHLMYGAKIDYKTIQVF 81
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 KRDHNYKSNRIVCYVGTWSVHKVDPTIEDIDPKCTHLMYGAKIDYKTIQVF 62
QY 82 DPYQDDNSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSDMAANPTYRQOFIQ 141
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63 DPYQDDNSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSDMAANPTYRQOFVQ 122
QY 142 SVLDFLQYKFDGLDWEYFGSRIGNPKIDKQNYLALVRELKDAFEPHGYLLTAASVPG 201
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 SVLDFLQYKFDGLDWEYFGSRIGNPKIDKQNYLALVRELKDAFEPHGYLLTAASVPG 182
QY 202 KDKIDRAVDIKELNKLFDWMNVMTYDYGHWENFGHNAPLYKRPDETDLHTYFNVNYT 261
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 KDKIDRAVDIKELNKLFDWMNVMTYDYGHWENFGHNAPLYKRPDETDLHTYFNVNYT 242
```

CC This protein is used in claimed fusion proteins of the invention. Such
 CC fusion proteins comprise a translocating moiety and a toxic moiety, where
 CC the translocating moiety is a plant protein (e.g. a lectin) capable of
 CC acting as a carrier to translocate the toxic moiety across the gut wall
 CC of a plant pathogen, and the toxic moiety is an insect-derived peptide or
 CC protein capable of causing deleterious effects on growth, development,
 CC reproduction or mortality in pest insects. Suitable insect peptides and
 CC proteins include allatostatins, chitinase, diuretic hormone and their
 CC metabolites and analogues. Polynucleotides encoding the fusion protein,
 CC vectors, host cells and transgenic plants that are resistant to disease
 CC are also provided. The fusion protein is target-specific, and resists
 CC degradation in the insect gut

XX Sequence 525 AA;

Query Match 32.5%; Score 1008.5; DB 6; Length 525;
 Best Local Similarity 38.4%; Pred. No. 7.7e-65;
 Matches 211; Conservative 93; Mismatches 187; Indels 59; Gaps 13;

QY 8 LSIACIGLNASIKRD-HNDYSKNPMRIVCYVGTWSVHKYD-PYTIEDIPFKCTHLM 65
 Db 5 VGVILVIAAFAAEHKAASAEKGVVYVGTWAVRPGNGRYDIEHIDFSLCTHLM 64
 QY 66 YGFAKIDVYKTIQVDPYQDDNHNHNSWEKRGYERFNNLRLKNPELTMISLGGWEGSGK 125
 Db 65 YGFFGINE-DATVRIIDPVLDEEN-WGRGHKRFVGLKNVGPGLTLAAIGWEGSGRK 122
 QY 126 YSDMAANPTVROQFIQSVLDFLQYKFDGLDWEYFGSLRGKPKIDKQNYLALVRELKD 185
 Db 123 FSAMAASGELRKEFISDCVAFQCRHGFQDGLDWEYPAQRDGNPLIDRDNHQAQVEEMRE 182
 QY 186 APFPHGYLTAASVPGOKIDRAYDIKELNKUFDWNNVNTYDHGWNENFYGHNAPLYKR 245
 Db 183 EFDHYGLLLTAASVAFSGVSYDIPRISKSFHFLNMVYDMHGAWDSYCGINAPLYRG 242
 QY 246 PDETDELTFYFNNTYTHYLLNNGATRDKLNVGVPYFGRAWSIDRESKLGDPKAGMSP 305
 Db 243 SADTDLRQGINVASIHFWLAQGGTGRKLVGLIPLYGRNFTLASAANTQIGAPTGGGT 302
 QY 306 PGFISEEGVLSYIELCOLFQKEEMHIOYDEVYNAFYGYNDKIWGYDDLASICKLAFL 365
 Db 303 VGYRTREPGVMGNECEKATEAMDLRSEQQVPIYARNQWGVYDILRSVQLKVKYL 362
 QY 366 KELGSGVMVWSLENDDFRGHC-GPKNPLLNKVNMINDEKNSPECILGPSTTTPTPT 424
 Db 363 LQGLGAMVWSLETDLFLGCGGGRYPLMHBSLVNGT-----PSTTTPSV 413
 QY 425 TPT 484
 Db 414 APPT-----STVAPGTTTTTTCANEGTTQPPT--SDAPNHTTSTTEGNGPTTRPSPG 466
 QY 485 TTEHTSETPKYTYVDGHLIKY--KEGDIHPHTNHHKYLVC-----EPVNGGWWH 535
 Db 467 -----DG---PCAGGRYGFVPHPTNCARYYICLTADTYEFT----- 500
 QY 536 IMPCPPGTIW 545
 Db 501 ---CPGTILF 507

RESULT 12

ID ABB58595
 XX ABB58595 standard; protein; 4498 AA.

AC ABB58595;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 2577.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

OS Drosophila melanogaster.

FN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL02698.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 2577; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (AB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 4498 AA;

Query Match 27.5%; Score 854.5; DB 4; Length 4498;
 Best Local Similarity 33.8%; Pred. No. 2e-52;
 Matches 202; Conservative 88; Mismatches 188; Indels 119; Gaps 21;

QY 2 KTIYALISVACIGLGNASIKRDHNDYSKNPMRIVCYVGTWSVYH-----KVDPTTIEDID 57
 Db 33 QTLFLCALAYCI-----NEASSEG-RVVCYVTNWSVYRPGTAKFNP---QNIN 77
 QY 58 PFKCTHLMY---GFAKIDVYKTIQVDPYQDDNHNHNSWEKRGYERFNNLRLKNPELTM 114
 Db 78 PYLCTHLYVAFGFTKDNQMK---PDKYQD-----IEQGYAKFTGLKTYNKQLKTM 128
 QY 115 SLGGWYEGSEKYSMAANPTVROQFIQSVLDFLQYKFDGLDWEYFGSLRGKPKIDKQ 174
 Db 129 AIGWNASRSPSLVASNERRQOFIKNLIKFLQNHFDGIDLDWEYPAHREGSKSRD 188
 QY 175 NYLALVRELKDAFEPHG-----YLLTAASVPGOKIDRAYDIKELNKUFDWNNVNTYD 227
 Db 189 NYAQVQELRAEFERAEKTRTELLTMAVPAGIEYDKGYDVPKLYLDWENLVLYD 248
 QY 228 YHGWENFYGHNAPLYKRPDETDELTY---FNNYTMHYLLNNGATRDKLNVGVPYGR 284
 Db 249 FHSSHEPSVNHAPLYSL--EEDSEVNYDAELNDYSIKYLLKAGADRDKLVLDIPTYGR 306
 QY 285 AWSIEDRSKLGDPKAGMSPPGFISGEEVLSYIELCOLFQKE-EWH-IQYDEVYNA 342
 Db 307 SYTLNEESTELGAPAGPGEGQGDATREKGYLAYEICQTLKDDPEVTVVQPNANVNGPY 366
 QY 343 GYNDKIWGVYDDLASISCKLAFLKELGVSGVWWSLENDDFRGHC-GPKNPLLNKVNMI 401
 Db 367 AYRNQWGVYDDDEALVVRKAEYVVAQGLGIMFWAINDDDFRGTCNCKPPLIEAA---- 422
 QY 402 NGDBKNSEFECILGPSTTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 445
 Db 423 ----KEAWVEALGLGINEVAKPSGPKPSRSRSDNASNRLNKGKTEAPLSSRRPSATR 478

| | | | |
|---|---|---|------------------------------------|
| Qy | 446 | SPPTPTTSPPTPTT | -----SPPTPTTPTPTAPTSTSTSPPTTTH 489 |
| | | : | : |
| Db | 479 | RPAVSTQAPP--PSTTFKLAEAGSSLYIGGRASTTTPPTTTPDP | -----522 |
| Qy | 490 | TSETPKTYTVDGHILKCYKGGDIPHTNIHKYLVCFVANG | -----GWWHIMPCPPG 542 |
| Db | 523 | -----GSDFKCEEFFQHRDCKKYTC--LDSGSLGLIVAHFTCPSG 566 | |
| RESULT 13 | | | |
| AAW01824 | | | |
| ID | AAW01824 standard; protein; 554 AA. | | |
| XX | | | |
| AC | AAW01824; | | |
| XX | | | |
| DT | 27-MAY-1997 (first entry) | | |
| XX | | | |
| DE | Manduca sexta larva epidermal and gut chitinase. | | |
| XX | | | |
| KW | Chitinase; insect; transgenic plant; biological control; baculovirus; | | |
| KW | biocide; insecticide; tobacco hookworm. | | |
| XX | | | |
| OS | Manduca sexta. | | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | Peptide | 1..19 | |
| FT | | /label= Sig_peptide | |
| FT | Protein | 20..554 | |
| FT | | /label= Mat_protein | |
| FT | Region | 97..111 | |
| FT | | /label= Conserved_region-I | |
| FT | Active-site | 136..148 | |
| FT | | /note= "proposed active site" | |
| XX | | | |
| PN | WQ9708944-A1. | | |
| XX | | | |
| PD | 13-MAR-1997. | | |
| XX | | | |
| PF | 06-SEP-1996; 96WO-US014440. | | |
| XX | | | |
| PR | 06-SEP-1995; 95US-00524051. | | |
| XX | | | |
| PA | (UNIV) UNIV KANSAS STATE RES FOUND. | | |
| XX | | | |
| PI | Kramer KJ, Muthukrishnan S, Choi HK, Corpuz L, Gopalakrishnan B; | | |
| XX | | | |
| DR | WPI; 1997-192586/17. | | |
| XX | | | |
| DR | N-PSDB; AAT62557. | | |
| XX | | | |
| PT | Transgenic plant contg. recombinant insect chitinase coding sequence - | | |
| PT | useful for controlling plant pests. | | |
| XX | | | |
| PS | Claim 4; Page 37-40; 57pp; English. | | |
| XX | | | |
| CC | An epidermal and gut chitinase (AAW01824) of Manduca sexta is encoded by | | |
| CC | cDNA clone 201 (AAT62557), isolated from a day-6 fifth instar whole larva | | |
| CC | cDNA library. The cDNA can be inserted into a vector for host expression | | |
| CC | of the chitinase protein, and may be utilised to control a population of | | |
| CC | insect pests. A recombinant baculovirus vector is provided for this | | |
| CC | purpose, as well as a transgenic tobacco plant that shows increased | | |
| CC | resistance to tobacco hornworm. Recombinant chitinase may also be | | |
| CC | expressed in host cells or larvae for subsequent purification and use as | | |
| CC | a biocide. The insect chitinase is expected to be more potent as an | | |
| CC | insect control agent than chitinases from other sources | | |
| XX | | | |
| SQ | Sequence 554 AA; | | |
| Query Match 27.0%; Score 838.5; DB 2; Length 554; | | | |
| Best Local Similarity 35.4%; Pred. No. 1.9e-52; | | | |
| Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 200 | | | |
| Qy | 6 | AIIUSWACTIGLNNASTIKRDHNDYSKNPMRIVCVGTWGVYHK-VDPYTIIDIPFKCTHL 64 | |
| | | : | : |

XX PS Disclosure; Col 53-58; 40pp; English.

XX CC The present sequence is the protein sequence for the Manduca sexta gut

CC chitinase. It was used to isolate the Drosophila melanogaster imaginal

CC disc growth factor 4 (IDGF4) coding sequence. IDGF4 is a member of the

CC chitinase related proteins (CHRP) and is involved in the promotion of

CC cell growth, motility and morphogenesis. The gene and protein are

CC expected to have mammalian homologues. They can be used in the treatment

CC of cancer, wound healing, tissue regeneration following arthritis,

CC osteoporosis, other skeletal disorders and burns, for revitalising scar

CC tissue resulting from surgical procedures, irradiation, laceration, toxic

CC chemicals, viral or bacterial infection or burns, to promote tissue

CC growth during tissue engineering, for example tissues for skin graft

CC replacements and bone regrowth, and to modulate the function of the

CC female reproductive tract. In addition, they can also be used to increase

CC meat, egg, sperm and milk production in animals. One possible method of

CC use is by gene therapy

XX SQ Sequence 554 AA;

Query Match 27.0%; Score 838.5; DB 3; Length 554;
Best Local Similarity 35.4%; Pred. No. 1.9e-52;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;

QY 6 AILSIMACIGLMNASIKRDNDYKPMRIYCVGTWVYHK-VDPYTIEDIDPKCTHL 64

Db 3 ATLATLAVLATAV-----QSDSRARIVCFNSWAVYRPGVGRYGIEDIPVEKCTHI 55

QY 65 MYGFAKIDEXYTIQVDFPDQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWYEGSE 124

Db 56 IYSGFIGVTEGSEVLIIDPELD-----VDKGRFNFTSLRSHSPSVKFWAVGWAEGSS 110

QY 125 KYSDMAANPTYRQQTQSVDLFDQYKFDGLDWEYFGSR-LGNPKIDKQNYLALVREL 183

Db 111 KYSHWAQKSTRMSFIRSVSFLKKYDFDGLDWEYFGAARDGGSFSDKDKFLYLVOEL 170

QY 184 KDAP--EPHGYLITAAVSPGKIDRAYDIKELKLPDMWVYTYDYGHWENFYGHNP 241

Db 171 RRARIRVKGWELTAAPLANFELMEGYHVPFELCOELDAIHVMSYDLRGNWAGFADVHSP 230

QY 242 LYKRPDETDLHYTFVNYTHYLLNNGATRDKLVMGVPYGRAWSIED-----290

Db 231 LYKRPED-QWAYEKLNVNDGLHLWEKGCPSNKLVLVGIPYGRSFTLSAGNNYGLGTI 289

QY 291 RSKLGLDPAKMGSPPGFISGEEGVLSYIELCOLFOKEE--WHIQYDEYVYAPYGNDKI 348

Db 290 NKEAGGDPAPYTNATG-----WAYVEICTEVDKDSGWTKKWDEQCKEYAYKGTQ 342

QY 349 WVGYYDLAS-CKLAFKELGVSGVMVWVSLNDDFKHCGKPNKPLNKVHNNMINGDEKNS 408

Db 343 WVGYPDPSVEIKMWNKQKYLCAWTAIDMDDFQGLCGEKPNLIKILHKMS-----396

QY 409 FEGILGSTTPTPTPTPTPTPT-----TPTPTS-----PPTPTTPTPTPTPTPTSP 456

Db 397 -----SYTVPPPHTEHTTPTPEWAPPPTSPDSEGDPIPTTTAKPASITTKTKTKT 449

QY 457 TPTPTPTPTPT-----PPT-TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 505

Db 450 TTTAKPPQSVDENDINVRPKPEPQPEPEVEVP-PTENB-----VDGSEI 497

QY 506 KCYKEGD-IHPHTNIHKYLVCFVNGVWVHIMPCPPGTTCQEKLTTC 552

Db 498 -CNSDQDIYDKKCDKYKWC--VNGE--AMQPSQCQGTGFVFNELNVC 540

RESULT 15
ABP72619
ID ABP72619 standard; protein; 554 AA.
XX AC ABP72619;
XX DT 11-JUN-2003 (first entry)

XX DE Manduca sexta chitinase.

XX KW Insecticide; pesticide; insect control; insect; toxin; chitinase; enzyme;

XX KW neuropeptide; transgenic plant; crop protection.

XX OS Manduca sexta.

XX FH Key Location/Qualifiers

FT Peptide 1..19

FT Protein /label= Signal_peptide 20..554

FT Modified-site /label= Mature_protein 85..88

FT Active-site /note= "Asn is N-glycosylated" 138..146

FT Modified-site 303..306

FT Modified-site /note= "Asn is N-glycosylated" 545..548

FT Modified-site /note= "Asn is N-glycosylated"

XX WO2003014150-A2.

XX PD 20-FEB-2003.

XX PF 06-AUG-2002; 2002WO-CB003598.

XX PR 08-AUG-2001; 2001GB-00019274.

XX PA (UYDU-) UNIV DURHAM.

XX PA (ENVI-) DEPT ENVIRONMENT FOOD & RURAL AFFAIRS.

XX PI Gatehouse JA, Fitches EC, Edwards JP;

XX WPI; 2003-278469/27.

XX PT Fusion protein useful for combating insect pests, comprises a

PT translocating moiety comprising a plant protein capable of acting as a

PT carrier to translocate toxic moiety inside plant pathogen, and a toxic

PT moiety.

XX Claim 7; Fig 10; 51pp; English.

XX CC The present sequence is that of Manduca sexta chitinase. This protein can

CC be used in claimed fusion proteins of the invention. Such fusion proteins

CC comprise a translocating moiety and a toxic moiety, where the

CC translocating moiety is a plant protein (e.g. a lectin) capable of acting

CC as a carrier to translocate the toxic moiety across the gut wall of a

CC plant pathogen, and the toxic moiety is an insect-derived peptide or

CC protein capable of causing deleterious effects on growth, development,

CC reproduction or mortality in pest insects. Suitable insect peptides and

CC proteins include allatostatin, chitinase, diuretic hormone and their

CC metabolites and analogues. Polynucleotides encoding the fusion protein,

CC vectors, host cells and transgenic plants that are resistant to disease

CC are also provided. The fusion protein is target-specific, and resists

CC degradation in the insect gut

XX SQ Sequence 554 AA;

Query Match 27.0%; Score 838.5; DB 6; Length 554;
Best Local Similarity 35.4%; Pred. No. 1.9e-52;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;

QY 6 AILSIMACIGLMNASIKRDNDYKPMRIYCVGTWVYHK-VDPYTIEDIDPKCTHL 64

Db 3 ATLATLAVLATAV-----QSDSRARIVCFNSWAVYRPGVGRYGIEDIPVEKCTHI 55

QY 65 MYGFAKIDEXYTIQVDFPDQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWYEGSE 124

Db 56 IYSGFIGVTEGSEVLIIDPELD-----VDKGRFNFTSLRSHSPSVKFWAVGWAEGSS 110

QY 125 KYSDMAANPTYRQQTQSVDLFDQYKFDGLDWEYFGSR-LGNPKIDKQNYLALVREL 183

Db 111 KYSHMVAQKSTRMSFIRSVVSLFKKYDFDGLDLDEWYPGAADRGSGFSKDKFYLVLVQEL 170
QY 184 KDAF--EPHGYLLTAAYSPGKIDRAYDIKELNKLFDWNNVNTYDHYHGWENFYGHNAP 241
Db 171 RRAFIRVGKWEITAAVPLANFRIMEGYHVPFELCQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDETDLHITYFNVNNTMYHLLNNGATRDKLVNGVPPFYGRAWSIED----- 290
Db 231 LYKRPDH-QWAYEKLNVNDGLHLWEKGCPSNKLVGIPFYGRSFTLSAGNNNYGLGTFI 289
QY 291 RSKLKLDPAKMSPPGFISGEGVLSYIELCQLFOKEE--WHIQYDEYNNAPYGYNDKI 348
Db 290 NKEAGGDEPAPYTNATG-----WAYYEICTEVDKDDSGWTKWDEQCKCFYAYKGTO 342
QY 349 WVGYYDDLASISCKLAFLKELGVSGVMVMSLENDDEPKHCGPKNPLANKVHNMINGDEKNS 408
Db 343 WVGYYEDPRSVEIKNNWIKQKGYLGAMTWAIDMDDFQGLCGERNPLIKILHKHVS----- 396
QY 409 FECILGPSTTPTPTPTPTPT-----TTPTPS-----PTTPTTPTPTPTPTTSP 456
Db 397 -----SYIVPPHTENTTPTPEWARPPSTPSDPSEGDPIPTTTAKPASITTKTVKTT 449
QY 457 TTPTTPTSPPT-----PTP-TTPTAPTSTPTSTPTTHTSETPKYTTYVDGHLI 505
Db 450 TTTTAKPPQSVIDEENDINVRPEPKPEPQPEPEVEVP-PTENE-----VDGSEI 497
QY 506 KCIYKED-IPHTNIHKYLVCEPVNGGWVHIMPCCPGTIWCOEKLTC 552
Db 498 -CNSDQDYIPDKKCDKYWRC--VNGE--AMQFSCQHGTVFNVELNVC 540

Search completed: March 22, 2004, 06:51:45
Job time : 170.752 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:59:34 ; Search time 113.53 Seconds
(without alignments)
1265.926 Million cell updates/sec

Title: US-09-662-293-15

Perfect score: 3107
Sequence: 1 MKTIYAILSIMACIGLNMNAS.....IMPCPPGTINQCKLTCIGE 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 3107 | 100.0 | 555 | 14 US-10-218-743-15 | Sequence 15, Appl |
| 2 | 3107 | 100.0 | 555 | 14 US-10-218-743-18 | Sequence 18, Appl |
| 3 | 3014 | 97.0 | 536 | 14 US-10-218-743-21 | Sequence 21, Appl |
| 4 | 2542 | 81.8 | 509 | 14 US-10-218-743-35 | Sequence 35, Appl |
| 5 | 2542 | 81.8 | 509 | 14 US-10-218-743-38 | Sequence 38, Appl |
| 6 | 2475 | 79.7 | 490 | 14 US-10-218-743-41 | Sequence 41, Appl |
| 7 | 738 | 23.8 | 445 | 14 US-10-004-219B-10 | Sequence 10, Appl |
| 8 | 738 | 23.8 | 466 | 14 US-10-161-547-2 | Sequence 2, Appl |
| 9 | 736.5 | 23.7 | 473 | 14 US-10-004-219B-4 | Sequence 4, Appl |
| 10 | 732 | 23.6 | 466 | 14 US-10-161-547-4 | Sequence 4, Appl |
| 11 | 729.5 | 23.5 | 452 | 14 US-10-004-219B-9 | Sequence 9, Appl |
| 12 | 721.5 | 23.2 | 373 | 14 US-10-161-547-14 | Sequence 14, Appl |
| 13 | 721.5 | 23.2 | 373 | 14 US-10-161-547-15 | Sequence 15, Appl |
| 14 | 708.5 | 22.8 | 455 | 14 US-10-004-219B-14 | Sequence 14, Appl |
| 15 | 708.5 | 22.8 | 476 | 14 US-10-004-219B-1 | Sequence 1, Appl |

| | | | | | |
|----|-------|------|-----|-------------------------|--------------------|
| 16 | 705.5 | 22.7 | 617 | 15 US-10-369-493-6743 | Sequence 6743, Ap |
| 17 | 668 | 21.5 | 678 | 15 US-10-295-027-632 | Sequence 632, App |
| 18 | 662.5 | 21.3 | 457 | 15 US-10-094-749-1900 | Sequence 1900, Ap |
| 19 | 662 | 21.3 | 383 | 14 US-10-097-340-45 | Sequence 45, Appl |
| 20 | 662 | 21.3 | 383 | 15 US-10-295-027-270 | Sequence 270, App |
| 21 | 650 | 20.9 | 383 | 9 US-09-459-749D-17 | Sequence 17, Appl |
| 22 | 630 | 20.3 | 385 | 15 US-10-295-027-278 | Sequence 278, App |
| 23 | 630 | 20.3 | 423 | 15 US-10-373-802-2 | Sequence 2, Appl |
| 24 | 592 | 19.1 | 599 | 15 US-10-369-493-17162 | Sequence 17162, A |
| 25 | 521.5 | 16.8 | 407 | 15 US-10-369-493-20271 | Sequence 20271, A |
| 26 | 516 | 16.6 | 428 | 15 US-10-369-493-16644 | Sequence 16644, A |
| 27 | 471 | 15.2 | 392 | 15 US-10-369-493-19361 | Sequence 19361, A |
| 28 | 458 | 14.7 | 351 | 15 US-10-369-493-12478 | Sequence 12478, A |
| 29 | 438 | 14.1 | 366 | 12 US-10-424-599-209667 | Sequence 209667, A |
| 30 | 414.5 | 13.3 | 485 | 15 US-10-369-493-3399 | Sequence 3399, Ap |
| 31 | 406 | 13.1 | 371 | 9 US-09-923-844B-2 | Sequence 2, Appl |
| 32 | 405 | 13.0 | 546 | 15 US-10-369-493-18584 | Sequence 18584, A |
| 33 | 401.5 | 12.9 | 399 | 14 US-10-316-754-11 | Sequence 11, Appl |
| 34 | 401.5 | 12.9 | 399 | 14 US-10-316-754-19 | Sequence 19, Appl |
| 35 | 372.5 | 12.0 | 449 | 15 US-10-369-493-3308 | Sequence 3308, Ap |
| 36 | 367.5 | 11.8 | 511 | 15 US-10-369-493-1684 | Sequence 1684, Ap |
| 37 | 352.5 | 11.3 | 572 | 15 US-10-369-493-18585 | Sequence 18585, A |
| 38 | 350.5 | 11.3 | 358 | 12 US-10-425-114-47084 | Sequence 47084, A |
| 39 | 350.5 | 11.3 | 365 | 12 US-10-424-599-273219 | Sequence 273219, A |
| 40 | 347.5 | 11.2 | 567 | 15 US-10-369-493-8158 | Sequence 8158, Ap |
| 41 | 337 | 10.8 | 789 | 15 US-10-369-493-12804 | Sequence 12804, A |
| 42 | 329.5 | 10.6 | 445 | 15 US-10-369-493-3089 | Sequence 3089, Ap |
| 43 | 317.5 | 10.2 | 452 | 15 US-10-369-493-8244 | Sequence 8244, Ap |
| 44 | 314 | 10.1 | 432 | 15 US-10-369-493-5521 | Sequence 5521, Ap |
| 45 | 313.5 | 10.1 | 607 | 14 US-10-156-761-10415 | Sequence 10415, A |

ALIGNMENTS

RESULT 1
US-10-218-743-15
; Sequence 15, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: 2002-08-13
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-15

Query Match 100.0%; Score 3107; DB 14; Length 555;
Best Local Similarity 100.0%; Pred. No. 3e-220;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKTIYAILSIMACIGLNMNASIKEDNDYSCNPMRIVCYVGTWSVYHKVDPTIEDIDPFK 60

Db 1 MKTIYAILSIMACIGLNMNASIKEDNDYSCNPMRIVCYVGTWSVYHKVDPTIEDIDPFK 60

QY 61 CTHLMYGFADKDEYKTIQVDFPDYQDDNHNHNSWKRGRYERFNNLRKNPELTTMISLGGWY 120
DB 61 CTHLMYGFADKDEYKTIQVDFPDYQDDNHNHNSWKRGRYERFNNLRKNPELTTMISLGGWY 120
QY 121 EGSEKYSDMAANPTYRQOFIQSVLDFLQEVKFDGLDWEYPGSRGPNKPKIDKQNYLALV 180
DB 121 EGSEKYSDMAANPTYRQOFIQSVLDFLQEVKFDGLDWEYPGSRGPNKPKIDKQNYLALV 180
QY 181 RELKDAFEPHGYLLTAAVSPGKOKIDRAYDIKELNKLFDWMNVTYDYHGGWENFYGHNA 240
DB 181 RELKDAFEPHGYLLTAAVSPGKOKIDRAYDIKELNKLFDWMNVTYDYHGGWENFYGHNA 240
QY 241 PLYKRPDETDELHTYFNWNTMYYLNGATRDKLWGVFPFYGRAMSIEDRSKLGDPDA 300
DB 241 PLYKRPDETDELHTYFNWNTMYYLNGATRDKLWGVFPFYGRAMSIEDRSKLGDPDA 300
QY 301 KGMSPPGFISGEGVLSYIELCOLFQKEEWHIOYDEYNAPYGYNDKIWGYDDLASISC 360
DB 301 KGMSPPGFISGEGVLSYIELCOLFQKEEWHIOYDEYNAPYGYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSGVMWVWLENDDFKHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
DB 361 KLAFLKELGVSGVMWVWLENDDFKHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
QY 421 TPSTTTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGGWWHIMPCP 540
DB 421 TPSTTTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGGWWHIMPCP 540
QY 541 PGTINCOEKLTCIGE 555
DB 541 PGTINCOEKLTCIGE 555

RESULT 2

US-10-218-743-18
; Sequence 18, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR FILING DATE: 2002-08-13
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-18

Query Match 100.0%; Score 3107; DB 14; Length 555;
Best Local Similarity 100.0%; Pred. No. 3e-220;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIIVAILSIACIGLMNASIKRDHNDYSKNPMRIVCVGTWVYHKVDPYTIIDIDPFFK 60
DB 1 MKTIIVAILSIACIGLMNASIKRDHNDYSKNPMRIVCVGTWVYHKVDPYTIIDIDPFFK 60
QY 61 CTHLMYGFADKDEYKTIQVDFPDYQDDNHNHNSWKRGRYERFNNLRKNPELTTMISLGGWY 120
DB 61 CTHLMYGFADKDEYKTIQVDFPDYQDDNHNHNSWKRGRYERFNNLRKNPELTTMISLGGWY 120
QY 121 EGSEKYSDMAANPTYRQOFIQSVLDFLQEVKFDGLDWEYPGSRGPNKPKIDKQNYLALV 180
DB 121 EGSEKYSDMAANPTYRQOFIQSVLDFLQEVKFDGLDWEYPGSRGPNKPKIDKQNYLALV 180
QY 181 RELKDAFEPHGYLLTAAVSPGKOKIDRAYDIKELNKLFDWMNVTYDYHGGWENFYGHNA 240
DB 181 RELKDAFEPHGYLLTAAVSPGKOKIDRAYDIKELNKLFDWMNVTYDYHGGWENFYGHNA 240
QY 241 PLYKRPDETDELHTYFNWNTMYYLNGATRDKLWGVFPFYGRAMSIEDRSKLGDPDA 300
DB 241 PLYKRPDETDELHTYFNWNTMYYLNGATRDKLWGVFPFYGRAMSIEDRSKLGDPDA 300
QY 301 KGMSPPGFISGEGVLSYIELCOLFQKEEWHIOYDEYNAPYGYNDKIWGYDDLASISC 360
DB 301 KGMSPPGFISGEGVLSYIELCOLFQKEEWHIOYDEYNAPYGYNDKIWGYDDLASISC 360
QY 361 KLAFLKELGVSGVMWVWLENDDFKHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
DB 361 KLAFLKELGVSGVMWVWLENDDFKHCQPKNPLLNKVNHNMGDEKNSFECILGPSTTTP 420
QY 421 TPSTTTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGGWWHIMPCP 540
DB 421 TPSTTTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGGWWHIMPCP 540
QY 541 PGTINCOEKLTCIGE 555
DB 541 PGTINCOEKLTCIGE 555

RESULT 3

US-10-218-743-21
; Sequence 21, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-21

Query Match 97.0%; Score 3014; DB 14; Length 536;

Best Local Similarity 100.0%; Pred. No. 2e-213;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SIKRDNDSKNPWRIVCYVGTWSVYHKVDPYTIEDIIDPKFCHLWYGAKEIDYKYTIQ 79
DB 1 SIKRDNDSKNPWRIVCYVGTWSVYHKVDPYTIEDIIDPKFCHLWYGAKEIDYKYTIQ 60
QY 80 VDPYQDDNHNSEKRGYVERFNNLRNLKLNPELTTMISLGWYEGSEKYSMAANPYRQOF 139
DB 61 VDPYQDDNHNSEKRGYVERFNNLRNLKLNPELTTMISLGWYEGSEKYSMAANPYRQOF 120
QY 140 IQSVLDFLOEYKFDGLDLDWEYPSGRSLGNPKIDQNYLALVRELKDAFPHGYLLTAAVS 199
DB 121 IQSVLDFLOEYKFDGLDLDWEYPSGRSLGNPKIDQNYLALVRELKDAFPHGYLLTAAVS 180
QY 200 PGKXIDRAYDIKELNKLFDMMNVTYDYHGCHENFYGHNAPLYKBPDETDLHTYFNVN 259
DB 181 PGKXIDRAYDIKELNKLFDMMNVTYDYHGCHENFYGHNAPLYKBPDETDLHTYFNVN 240
QY 260 YTMHYLLANGATRDKLWGVFPFYGRAWSIEDRSKLLGDPAGKMSPPGFI SGEEGVLSYI 319
DB 241 YTMHYLLANGATRDKLWGVFPFYGRAWSIEDRSKLLGDPAGKMSPPGFI SGEEGVLSYI 300
QY 320 ELCQFQKEEWHIOYDEYNAAPYGNDKI WGVYDDLASISCKLAFELKELGVSGVMVWSLE 379
DB 301 ELCQFQKEEWHIOYDEYNAAPYGNDKI WGVYDDLASISCKLAFELKELGVSGVMVWSLE 360
QY 380 NDDFKHGCGKPNLLKNVHNMINGDEKNSPECILGPSTTTPTPTTPTTPTTPTTPTTPT 439
DB 361 NDDFKHGCGKPNLLKNVHNMINGDEKNSPECILGPSTTTPTPTTPTTPTTPTTPTTPT 420
QY 440 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
DB 421 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
QY 500 VDGHLIKYKCGDIPHTPHNIHKYLVCFVNGVWVHIMPCPGTIWCQEKLCIGE 555
DB 481 VDGHLIKYKCGDIPHTPHNIHKYLVCFVNGVWVHIMPCPGTIWCQEKLCIGE 536

RESULT 4
US-10-218-743-35
; Sequence 35, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-35

Query Match 81.8%; Score 3542; DB 14; Length 509;
Best Local Similarity 81.7%; Pred. No. 1e-178;

Query Match 81.8%; Score 2542; DB 14; Length 509;
Best Local Similarity 81.7%; Pred. No. 1e-178;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

QY 1 MKTITAILSIMACIGLNASIKRDHNDYSKNPMEIIVCYGTWVYHKVDPTTIEDIDPFK 60
DB 1 MKTIFALFCIWACIGLWNAATKRDHNNYSKNPMEIIVCYGTWVYHKVDPTTIEDIDPFK 60

QY 61 CTHLMYGFADIKYKTIQVDFPYDDNNHNSWKGVERFNNLRKNPELTMTISLGGWY 120
DB 61 CTHLMYGFADIKYKTIQVDFPYDDNNHNSWKGVERFNNLRKNPELTMTISLGGWY 120

QY 121 EGSEKYSMDAANPYROOFTQSVLDFLOEYKFGDLDDWEYFGSRLGNPKIDKQNYLALV 180
DB 121 EGSEKYSMDAANPYROOFTQSVLDFLOEYKFGDLDDWEYFGSRLGNPKIDKQNYLALV 180

QY 181 RELKDAEPFPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
DB 181 RELKDAEPFPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240

QY 241 PLYKRPDETDELHTYFNNVMTYLLNGATRDCLVGVFPFYGRAMSIEDRSKVLGDDPA 300
DB 241 PLYKRPDETDELHTYFNNVMTYLLNGATRDCLVGVFPFYGRAMSIEDRSKVLGDDPA 300

QY 301 KGMSPPGFISGEGVLSYIELCOLFOKEEWHIQYDEYNAPYGYNDKIWVGYDDLASISC 360
DB 301 KGMSPPGFISGEGVLSYIELCOLFOKEEWHIQYDEYNAPYGYNDKIWVGYDDLASISC 360

QY 361 KLAFLKELGSGVVMWSLENDDFKHCCKPKNPLNKVHNNINGDEKNSFECILGSPSTTP 420
DB 361 KLAFLKELGSGVVMWSLENDDFKHCCKPKNPLNKVHNNINGDEKNSFECILGSPSTTP 420

QY 421 TPPT 480
DB 421 TPPT 480

QY 481 TPSTTTTHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKLVCEV---NGSWVWHIM 537
DB 435 ---PTTDTSTSETPKYTYIDGHLIKCYKQGLPHPTDVHKLVCEVIATENGWVWHIM 491

QY 538 PCPFGTWCQKLCIGCE 555
DB 492 DCPKGTWHTATLKNCIQE 509

RESULT 6
US-10-218-743-41
; Sequence 41, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490

TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-41

Query Match 79.7%; Score 2475; DB 14; Length 490;
Best Local Similarity 82.5%; Pred. No. 8.4e-174;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;

QY 22 KRDHNDYSKNPMEIIVCYGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQVF 81
DB 3 KRDHNNYSKNPMEIIVCYGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQVF 62

QY 82 DPYQDDNNSWKGVERFNNLRKNPELTMTISLGGWYEGSEKYSMDAANPYRQOQFIQ 141
DB 63 DPYQDDNNSWKGVERFNNLRKNPELTMTISLGGWYEGSEKYSMDAANPYRQOQFVQ 122

QY 142 SVLDFLOEYKFGDLDDWEYFGSRLGNPKIDKQNYLALVRELKDAEPFPHGYLLTAAVSPG 201
DB 123 SVLDFLOEYKFGDLDDWEYFGSRLGNPKIDKQNYLALVRELKDAEPFPHGYLLTAAVSPG 182

QY 202 KDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNNVMT 261
DB 183 KDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNNVMT 242

QY 262 MEYLLNNGATRDCLVGVFPFYGRAMSIEDRSKVLGDDPAKGMSPPGFISGEGVLSYIEL 321
DB 243 MEYLLNNGATRDCLVGVFPFYGRAMSIEDRSKVLGDDPAKGMSPPGFISGEGVLSYIEL 302

QY 322 COLFOKEEWHIQYDEYNAPYGYNDKIWVGYDDLASISCKLAFKELGSGVVMWSLEND 381
DB 303 COLFOKEEWHIQYDEYNAPYGYNDKIWVGYDDLASISCKLAFKELGSGVVMWSLEND 362

QY 382 DFKHCGCKPKNPLNKVHNNINGDEKNSFECILGSPSTPTPTPTPTPTPTPTPTPTPT 441
DB 363 DFKHCGCKPKNPLNKVHNNINGDEKNSFECILGSPSTPTPTPTPTPTPTPTPTPTPT 415

QY 442 TPT 501
DB 416 -----PTTDTSTSETPKYTYID 433

QY 502 GHLIKCYKEGDIHPHTNIHKLVCEV---NGSWVWHIMPCPFGTWCQKLCIGCE 555
DB 434 GHLIKCYKQGLPHPTDVHKLVCEVIATENGWVWHIMDCPKGTWHTATLKNCIQE 490

RESULT 7
US-10-004-2198-10
; Sequence 10, Application US/100042198
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,2198
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(445)
; OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-2198-10

Query Match 23.8%; Score 738; DB 14; Length 445;
Best Local Similarity 34.2%; Pred. No. 5.3e-46;

Db 292 GAYTRQAGWAYEICTFLRSGATEVM-----DASQEVFPYAYKANWLGVDYNIKSFSVKAQ 347
QY 364 FLKELGVSVMWSLENDKFG-HCG-----PKNPLLNKVNHMINGDEKNSFECILGPSTT 418
Db 348 WLKQNFEGGAMWALDUDFTGSCDQKFPPLSTLTKA-----LGISTE 392
QY 419 TPTPTTPTTPTTPTT 433
Db 393 GCTAPDVPSEVPVTP 407

RESULT 10
US-10-161-547-4
; Sequence 4, Application US/10161547
; Publication No. US2003014216A1
; GENERAL INFORMATION:
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1) (452)
; OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

Query Match 23.5%; Score 729.5; DB 14; Length 452;
Best Local Similarity 37.0%; Pred. No. 2.3e-45;
Matches 154; Conservative 69; Mismatches 144; Indels 49; Gaps 11;

QY 35 IVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGFADKIDYKTIQVDFPDYQDDNHSW- 92
Db 3 LICFTNNAQYRPGIGSKPDDINPCLCTHLYAFAGMNNETIT-----IENW 51
QY 93 EKRGYERFNNLRKNPETLTMISLGGWYEGSEKYSDMAANFYRQOFIOSVLDLFOEYK 152
Db 52 DVTLYKAFNDLKNRNSKLTLLAIGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGF 111
QY 153 DGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASVSGDKI 205
Db 112 DGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASVSGDKI 170
QY 206 DRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDTEDELHTYFNVNVTMYY 265
Db 171 QAGYIPELSKYLDPIHVTYDLHGSWEGYTGENSEPLYKYPTETGS-NAYLNVYVWY 229
QY 266 LNNGATRDKLVMGVPPYGRAMSIEDRSKLGDPAGKMSPPGFISGERGVLSYIELCOLF 325
Db 230 KNGGAPAEKLVIGFPEYGHFTFLRNPDSNGIGAPTSGDGPAGAYTRQAGFWAYEICTFL 289
QY 326 QK---BEWHIQVDEYNAPYGVNDKIWGYDDLASISCKLAFKELGVSVMWSLEND 382
Db 290 RSGATEVM-----DASQEVFPYAYKANWLGVDYNIKSFSVKAQWLNKQNFEGGAMWALD 345
QY 383 FKG-HCG-----PKNPLLNKVNHMINGDEKNSPECILGPSTTPTTPTTPTTPTT 433
Db 346 FTGSCDQKFPPLSTLTKA-----LGISTEACTAPDVPSEVPVTP 386

RESULT 12
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39

Db 292 GAYTRQAGWAYEICTFLRSGATEVM-----DASQEVFPYAYKANWLGVDYNIKSFSVKAQ 347
QY 364 FLKELGVSVMWSLENDKFG-HCG-----PKNPLLNKVNHMINGDEKNSFECILGPSTT 418
Db 348 WLKQNFEGGAMWALDUDFTGSCDQKFPPLSTLTKA-----LGISTE 392
QY 419 TPTPTTPTTPTTPTT 433
Db 393 GCTAPDVPSEVPVTP 407

RESULT 10
US-10-161-547-4
; Sequence 4, Application US/10161547
; Publication No. US2003014216A1
; GENERAL INFORMATION:
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1) (452)
; OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

Query Match 23.5%; Score 729.5; DB 14; Length 452;
Best Local Similarity 37.0%; Pred. No. 2.3e-45;
Matches 154; Conservative 69; Mismatches 144; Indels 49; Gaps 11;

QY 35 IVCYVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGFADKIDYKTIQVDFPDYQDDNHSW- 92
Db 3 LICFTNNAQYRPGIGSKPDDINPCLCTHLYAFAGMNNETIT-----IENW 51
QY 93 EKRGYERFNNLRKNPETLTMISLGGWYEGSEKYSDMAANFYRQOFIOSVLDLFOEYK 152
Db 52 DVTLYKAFNDLKNRNSKLTLLAIGWNFGTAPFTTWSTQNRQTFITSVIKFLROYGF 111
QY 153 DGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASVSGDKI 205
Db 112 DGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAASVSGDKI 170
QY 206 DRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDTEDELHTYFNVNVTMYY 265
Db 171 QAGYIPELSKYLDPIHVTYDLHGSWEGYTGENSEPLYKYPTETGS-NAYLNVYVWY 229
QY 266 LNNGATRDKLVMGVPPYGRAMSIEDRSKLGDPAGKMSPPGFISGERGVLSYIELCOLF 325
Db 230 KNGGAPAEKLVIGFPEYGHFTFLRNPDSNGIGAPTSGDGPAGAYTRQAGFWAYEICTFL 289
QY 326 QK---BEWHIQVDEYNAPYGVNDKIWGYDDLASISCKLAFKELGVSVMWSLEND 382
Db 290 RSGATEVM-----DASQEVFPYAYKANWLGVDYNIKSFSVKAQWLNKQNFEGGAMWALD 345
QY 383 FKG-HCG-----PKNPLLNKVNHMINGDEKNSPECILGPSTTPTTPTTPTTPTT 433
Db 346 FTGSCDQKFPPLSTLTKA-----LGISTEACTAPDVPSEVPVTP 386

RESULT 12
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Tjoelker, Larry W.
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-14

Query Match      23.2%; Score 721.5; DB 14; Length 373;
Best Local Similarity 37.7%; Pred. No. 6.9e-45;
Matches 136; Conservative 83; Mismatches 117; Indels 25; Gaps 8;

QY 34 RIVCYVGTWSYHKVDP--YTIEDIDPFKCTHLMYGFAKIDEXKYTIQVDFPYQDDNHSW 92
Db 2 KLVICYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGTNNHLSLST-----TEW 50

QY 93 -EKRGYERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANTYQQFIQSVDLFQBYK 151
Db 51 NDETLYOEFNGLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYS 110

QY 152 FDGLDLDWEYPGSRGRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAAVSPGKDK 204
Db 111 FDGLDLDWEYPGSQ--GSPAVDKERFTTLVQDLANAFQQAQTSQGERLLLSAAVPAQOTY 169

QY 205 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKRPDDETLHTYFNNVNTMHY 264
Db 170 VDAGYEVDKIAQNLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAAS-LNVDAAVQO 228

QY 265 YLNNGATRDKLVGMVFPYGRAWSIEDRSKLGDPKAGMSPPGFTSGEEGLVSYIELCQL 324
Db 229 WLQKGTTPASKLILGMPTYGRSFTLASSSDTRVGAPATSGTGPFTKEGGLMAYEVCWS 288

QY 325 FQKEWHIQDEYNNAPYGYNDKIWGYDDIASISCKLAFKELGVSGVMWVSLNDPDK 384
Db 289 KGATKQRIQ--DQ--KVPYIFRDNQWGFDDVESFKTKVSYLKQKGLGGAMVWALDLDFA 345

QY 385 G 385
Db 346 G 346

RESULT 13
US-10-161-547-15
; Sequence 15, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-15

Query Match      23.2%; Score 721.5; DB 14; Length 373;
Best Local Similarity 37.7%; Pred. No. 6.9e-45;
Matches 136; Conservative 83; Mismatches 117; Indels 25; Gaps 8;

QY 34 RIVCYVGTWSYHKVDP--YTIEDIDPFKCTHLMYGFAKIDEXKYTIQVDFPYQDDNHSW 92
Db 2 KLVICYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGTNNHLSLST-----TEW 50

QY 93 -EKRGYERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANTYQQFIQSVDLFQBYK 151
Db 51 NDETLYOEFNGLKKNPKLKTLLAIGGNFQTKFTDMVATANNRQTFVNSAIRFLRKYS 110

QY 152 FDGLDLDWEYPGSRGRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAAVSPGKDK 204
Db 111 FDGLDLDWEYPGSQ--GSPAVDKERFTTLVQDLANAFQQAQTSQGERLLLSAAVPAQOTY 169

QY 205 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKRPDDETLHTYFNNVNTMHY 264
Db 170 VDAGYEVDKIAQNLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAAS-LNVDAAVQO 228

QY 265 YLNNGATRDKLVGMVFPYGRAWSIEDRSKLGDPKAGMSPPGFTSGEEGLVSYIELCQL 324
Db 229 WLQKGTTPASKLILGMPTYGRSFTLASSSDTRVGAPATSGTGPFTKEGGLMAYEVCWS 288

QY 325 FQKEWHIQDEYNNAPYGYNDKIWGYDDIASISCKLAFKELGVSGVMWVSLNDPDK 384
Db 289 KGATKQRIQ--DQ--KVPYIFRDNQWGFDDVESFKTKVSYLKQKGLGGAMVWALDLDFA 345

QY 385 G 385
Db 346 G 346

RESULT 14
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boet, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)-(455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14

Query Match      22.8%; Score 708.5; DB 14; Length 455;
Best Local Similarity 35.9%; Pred. No. 8.1e-44;
Matches 149; Conservative 75; Mismatches 154; Indels 37; Gaps 10;

QY 34 RIVCYVGTWSYHK--VDPYTIEDIDPFKCTHLMYGFAKIDEXKYTIQVDFPYQDDNHSW 92
Db 2 QLTCTVTNWAQYRGLGRFMDNIDPCLCTHLIYAFAGRNNEITT-----IEW 50

QY 93 -EKRGYERFNNLRKKNPELTMTISLGGWYEGSEKYSDMAANTYQQFIQSVDLFQBYK 151
Db 51 NDVTLYOAFNGLKKNKNSQLTKLLAIGGNFQTAFTAMVSTPENRQTFITSVIKFLRQE 110

QY 152 FDGLDLDWEYPGSRGRLGNPKIDKQNYLALVRELKDAPEPHG-----YLLTAAVSPGKDK 204
Db 111 FDGLDLDWEYPGSR--GSPQDQKHLFTVLVQEMREAFQEAQKINKPLMWTAAVAGISN 169

QY 205 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKRPDDETLHTYFNNVNTMHY 264
Db 170 IQSGYEIPLQSQYLDYIHVMYDLHGSWEYTGENSELYKYPTDGTGS-NAYLVNDYVMNY 228

QY 265 YLNNGATRDKLVGMVFPYGRAWSIEDRSKLGDPKAGMSPPGFTSGEEGLVSYIELCQL 324
Db 229 WIDNGAPAEKLVIGFPYTYGHNFILSNFTSGAPTSGAGPAGPYAKESGIWAYYEIC-T 287
```


Db 117 FDGLDMDEYPTQGGSPD-DYDNFVILMAELNQAHAEGMLLTAASAGKATIDPAYNV 175
QY 212 KEINKLFDWNTVTDYHGGWENFYCHNAPLYKRP-DETDLHTYFNVNYTMEYYLNGA 270
Db 176 PELSKELDLINWTDYLRGAWDDYTHQSGLYAHLDEGD--NAFLNVDFAISYWIEKGA 233
QY 271 TRDKLVWGVPFYGRAWSIEDRSKLGDPAGKMSPPGFISGEEGVLSELCQLFQKEEW 330
Db 234 RGOIALGIFLYGRCWTLASQQETGYYPAHQPGAAGDWTRSPGMLGYNEICYMRTTQDW 293
QY 331 HIQYDEYYNAPGY---NDKIWVGDDLASISCKLAFKELGVSGVWWSLENDDEKHC 387
Db 294 TVVDDPAMNEPYTYFPNMNINWCSYDHAASVAIKAEYAKSKGLAGTMWWSVETDDFRGLC 353
QY 388 GPKNPLINKVHNMINGDEKNSPECILGPTTPTTPTTPTTPTTPTTPTTPTTPTTSP 447
Db 354 H-----NRKYHLI-----KTMVEVFGGSITEP----- 376
QY 448 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 507
Db 377 -----PPLPTTTRDPNEFTTTTRAPP-----PG-----VHC 403
QY 508 YKEGDIPHPTNIHKYLVCEF-VNGGWVHIMPCPPGTIWCQEKLT 552
Db 404 TQGLNPDPLDCTHYILCSLNTSGYNEKEVEVCEGTLYNPQSYC 449

Search completed: March 22, 2004, 06:59:26
Job time : 113.679 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:30:23 ; Search time 168.752 Seconds
(without alignments)
929.256 Million cell updates/sec

Title: US-09-662-293-18
Perfect score: 3107
Sequence: 1 MKTYAIISSIMACIGIMNAS.....IMPCPPGTTCQKLTICGE 555

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|---------------------|
| 1 | 3107 | 100.0 | 555 | AAV52523 | House dus |
| 2 | 3107 | 100.0 | 555 | AAU96327 | Der HMW-m |
| 3 | 3107 | 100.0 | 555 | AAU96328 | Der HMW-m |
| 4 | 3014 | 97.0 | 536 | AAV52525 | House dus |
| 5 | 3014 | 97.0 | 536 | AAU96329 | Der HMW-m |
| 6 | 2542 | 81.8 | 509 | AAV52533 | D. pteron |
| 7 | 2542 | 81.8 | 509 | AAU96337 | Der HMW-m |
| 8 | 2542 | 81.8 | 509 | AAU96338 | Der HMW-m |
| 9 | 2475 | 79.7 | 490 | AAV52535 | D. pteron |
| 10 | 2475 | 79.7 | 490 | AAU96339 | Der HMW-m |
| 11 | 1008.5 | 32.5 | 525 | ABP72636 | Abp72636 Anopheles |
| 12 | 854.5 | 27.5 | 498 | ABP72634 | Abp72634 Anopheles |
| 13 | 838.5 | 27.0 | 554 | AAW01824 | Manduca s |
| 14 | 838.5 | 27.0 | 554 | AAW01824 | Manduca s |
| 15 | 838.5 | 27.0 | 554 | AAW01824 | Manduca s |
| 16 | 837.5 | 27.0 | 467 | ABP72619 | Abp72619 Manduca s |
| 17 | 837.5 | 27.0 | 565 | ABP72634 | Abp72634 Prayn chi |
| 18 | 837.5 | 27.0 | 595 | ABP72625 | Abp72625 Bombyx mo |
| 19 | 832 | 26.8 | 460 | ABP72636 | Abp72636 Drosophila |
| 20 | 815.5 | 26.2 | 583 | AAE28197 | AAE28197 Flea chit |
| 21 | 812.5 | 26.2 | 635 | AAE28203 | AAE28203 Flea chit |
| 22 | 811.5 | 26.1 | 559 | AAE28199 | AAE28199 Flea pfc |
| 23 | 810 | 26.1 | 574 | ABP72635 | Abp72635 Aedes aeg |
| 24 | 802 | 25.8 | 483 | ABP72633 | Abp72633 Chelonus |
| 25 | 787 | 25.3 | 553 | ABP72626 | Abp72626 Hyphantri |

| | | | | | | |
|----|-------|------|-----|---|----------|-----------|
| 26 | 738 | 23.8 | 466 | 2 | AAW08584 | Human 50 |
| 27 | 738 | 23.8 | 466 | 2 | AAW40259 | Human chi |
| 28 | 738 | 23.8 | 466 | 2 | AAW42425 | MO-218 cl |
| 29 | 738 | 23.8 | 466 | 4 | AAE00432 | Human chi |
| 30 | 738 | 23.8 | 466 | 5 | AAE25903 | Human chi |
| 31 | 738 | 23.8 | 466 | 5 | ABB76291 | Human chi |
| 32 | 737.5 | 23.7 | 520 | 6 | ABU09914 | Partial m |
| 33 | 736.5 | 23.7 | 473 | 7 | ABR55543 | Amino aci |
| 34 | 735.5 | 23.7 | 459 | 7 | ADC24231 | Human MOV |
| 35 | 734.5 | 23.6 | 473 | 7 | ADC51464 | Chitotria |
| 36 | 732.5 | 23.6 | 387 | 2 | AAW08585 | Human 39 |
| 37 | 732 | 23.6 | 447 | 7 | ADC24237 | Human NOV |
| 38 | 732 | 23.6 | 466 | 2 | AAW40260 | Human chi |
| 39 | 732 | 23.6 | 466 | 2 | AAW42426 | MO-13B cl |
| 40 | 732 | 23.6 | 466 | 4 | AAE00433 | Human chi |
| 41 | 732 | 23.6 | 466 | 5 | AAE25904 | Human chi |
| 42 | 732 | 23.6 | 466 | 5 | ABB76292 | Human chi |
| 43 | 731 | 23.5 | 466 | 2 | AAW31498 | Human chi |
| 44 | 729.5 | 23.5 | 452 | 7 | ABR55544 | Amino aci |
| 45 | 723 | 23.3 | 466 | 6 | ABP72621 | Human chi |

ALIGNMENTS

RESULT 1
AAV52523

| | | |
|----|---|--|
| ID | AAV52523 | standard; protein; 555 AA. |
| XX | AC | AAV52523; |
| XX | AC | AAV52523; |
| DT | 22-FEB-2000 | (first entry) |
| XX | House dust mite (D. farinae) | mite allergen protein (map) PDerf98-555. |
| XX | Mite allergen protein; map; | high molecular weight; HMW-map; allergy; |
| XX | house dust mite; IGE; | immunoglobulin E; allergen; mapA; mapB; |
| XX | hypersensitivity reaction; therapy; | treatment; diagnosis; human; feline; |
| XX | canine; veterinary; antibody; | vaccine; immunisation. |
| XX | Dermatophagoides farinae. | |
| XX | Key | Location/Qualifiers |
| FT | Peptide | 1..19 |
| FT | Protein | /note= "Signal peptide" |
| FT | Protein | 20..555 |
| FT | Protein | /note= "Mature Pderf98-555" |
| XX | WO9954349-A2. | |
| XX | 28-OCT-1999. | |
| XX | 16-APR-1999; | 99WO-US0008524. |
| XX | 17-APR-1998; | 98US-00062013. |
| XX | 13-MAY-1998; | 98US-0085295P. |
| XX | 02-SEP-1998; | 98US-0098909P. |
| XX | (HESK-) HESKA CORP. | |
| XX | Mccall CA, Hunter SW, Weber ER; | |
| XX | WPI; 2000-052700/04. | |
| XX | N-PSDB; AAZ38575, AAZ38576, AAZ38577, AAZ38578. | |
| XX | Novel high molecular weight Dermatophagoides nucleic acid polypeptides | |
| XX | used to modify an animals' hypersensitivity to mite allergens. | |
| XX | Claim 3; Page 111-113; 154pp; English. | |
| XX | This sequence represents Dermatophagoides farinae mite allergen protein | |
| XX | (map) PDerf98-555. Pderf98-555 has a molecular weight of 98 kD, | |
| XX | comprising 555 amino acids, and is a component of the Dermatophagoides | |

CC farinae high molecular weight mite allergen protein (HMW-map)
CC composition. The HMW-map composition was isolated from a D. farinae
CC homogenate by gel filtration, with each fraction being analysed for the
CC presence of proteins that bound to IgE present in mite-allergic dog
CC antisera. Mite allergenic proteins and peptides, and nucleic acids
CC encoding them, may be used in therapeutic compositions to modify an
CC animal's hypersensitivity reaction to mite allergens. Animals that may be
CC treated include mammals and birds, especially felines, canines, equines,
CC humans, other pets, and work or domestic animals. The proteins or
CC fragments may also be used to diagnose allergies via a skin test. The
CC proteins and peptides can also be used to raise antibodies, which have a
CC variety of potential uses. For example, they can be used as vaccines to
CC passively immunise animals against dust mite hypersensitivity, as
CC positive controls in test kits and as tools to recover desired dust mite
CC allergens from a mixture of proteins
XX
SQ Sequence 555 AA;

Query Match 100.0%; Score 3107; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.8e-217;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWVSVYHKVDPTTIEDIDPFK 60
DB 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWVSVYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGFADIEYKTYTQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWY 120
DB 61 CTHLMYGFADIEYKTYTQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWY 120
QY 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGLGNPKIDQNYLALV 180
DB 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGLGNPKIDQNYLALV 180
QY 181 RELKDAFEHGYLLTAASVPGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
DB 181 RELKDAFEHGYLLTAASVPGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
QY 241 PLYKRPDETDLHTYFNVNNTMYLLNNGATRDKLVMGVFFYGRANSIEDRSKLGDP 300
DB 241 PLYKRPDETDLHTYFNVNNTMYLLNNGATRDKLVMGVFFYGRANSIEDRSKLGDP 300
QY 301 KGMSPPGFISGEEGVLSIELCQLFQKEWHIQVDEYNNAPYGYNDKIWVGDDLASISC 360
DB 301 KGMSPPGFISGEEGVLSIELCQLFQKEWHIQVDEYNNAPYGYNDKIWVGDDLASISC 360
QY 361 KLAFLKELGVSGVMVWSLENDDFKHCPCNPLNKHVNNINGDEKNSFECILGPSTTP 420
DB 361 KLAFLKELGVSGVMVWSLENDDFKHCPCNPLNKHVNNINGDEKNSFECILGPSTTP 420
QY 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
DB 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
QY 481 TSPPTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLCFVNGWVHIMPCP 540
DB 481 TSPPTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLCFVNGWVHIMPCP 540
QY 541 PGTIWCQEKLTICGE 555
DB 541 PGTIWCQEKLTICGE 555

RESULT 2

AAU96327
ID AAU96327 standard; protein; 555 AA.

XX
AC AAU96327;

XX
DT 15-JUL-2002 (first entry)

XX
DE Der HMW-map polypeptide #14.

KW Der HMW-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
OS Dermatophagoides farinae.
PN WO200222807-A2.
PD 21-MAR-2002.
PF 14-SEP-2001; 2001WO-US028730.
XX 14-SEP-2000; 2000US-00662293.
PR (HESK-) HESKA CORP.
PA
XX
PI McCall CA, Hunter SW, Weber ER;
XX
DR WPI; 2002-351888/38.
DR N-PSDB; ABK69571.
XX
PT New mite allergenic protein isolated from Dermatophagoides, designated
PT Der HMW-map protein, useful as a vaccine for treating mite allergy.
XX
PS Claim 12; Page 114-116; 161pp; English.
XX

CC The invention relates to an isolated mite allergenic protein of
CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
CC acid. The Der HMW-map protein is useful for eliciting an immune response
CC against Der HMW-map protein. The protein or a reagent comprising a non-
CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
CC cat) susceptible to or having an allergic response to a mite. A
CC therapeutic composition is useful for desensitising a host animal to an
CC allergic response to a mite. The DNA and protein can be used in the
CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
CC binding of proteins to IgE, to prevent immunocomplex formation, thus
CC reducing hypersensitivity responses to mite allergens, and as vaccines
CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
CC represent Der HMW-map polypeptides of the invention
XX

SQ Sequence 555 AA;

Query Match 100.0%; Score 3107; DB 5; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.8e-217;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWVSVYHKVDPTTIEDIDPFK 60
DB 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWVSVYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGFADIEYKTYTQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWY 120
DB 61 CTHLMYGFADIEYKTYTQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGWY 120
QY 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGLGNPKIDQNYLALV 180
DB 121 EGSEKYSDMAANPYRQOFTQSVLDFLOEYKFDGLDWEYPSGLGNPKIDQNYLALV 180
QY 181 RELKDAFEHGYLLTAASVPGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
DB 181 RELKDAFEHGYLLTAASVPGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
QY 241 PLYKRPDETDLHTYFNVNNTMYLLNNGATRDKLVMGVFFYGRANSIEDRSKLGDP 300
DB 241 PLYKRPDETDLHTYFNVNNTMYLLNNGATRDKLVMGVFFYGRANSIEDRSKLGDP 300
QY 301 KGMSPPGFISGEEGVLSIELCQLFQKEWHIQVDEYNNAPYGYNDKIWVGDDLASISC 360
DB 301 KGMSPPGFISGEEGVLSIELCQLFQKEWHIQVDEYNNAPYGYNDKIWVGDDLASISC 360
QY 361 KLAFLKELGVSGVMVWSLENDDFKHCPCNPLNKHVNNINGDEKNSFECILGPSTTP 420


```

DR WPI; 2000-052700/04.
XX N-PSDB; AA238579, AA238580.
PT Novel high molecular weight Dermatophagoides nucleic acid polypeptides
PT used to modify an animals' hypersensitivity to mite allergens.
XX PS Claim 3; Page 125-127; 154pp; English.
XX This sequence represents Dermatophagoides farinae mite allergen protein
CC (map) pDerf98-536, the mature form of pDerf98-555 (AAV52523). pDerf98-536
CC has a molecular weight of 98 kD, comprising 536 amino acids, and is a
CC component of the Dermatophagoides farinae high molecular weight mite
CC allergen protein (HWM-map) composition. The HWM-map composition was
CC isolated from a D. farinae homogenate by gel filtration, with each
CC fraction being analysed for the presence of proteins that bound to IgE
CC present in mite-allergic dog antisera. Mite allergenic proteins and
CC peptides, and nucleic acids encoding them, may be used in therapeutic
CC compositions to modify an animal's hypersensitivity reaction to mite
CC allergens. Animals that may be treated include mammals and birds,
CC especially felines, canines, equines, humans, other pets, and work or
CC domestic animals. The proteins or fragments may also be used to diagnose
CC allergies via a skin test. The proteins and peptides can also be used to
CC raise antibodies, which have a variety of potential uses. For example,
CC they can be used as vaccines to passively immunise animals against dust
CC mite hypersensitivity, as positive controls in test kits and as tools to
CC recover desired dust mite allergens from a mixture of proteins
XX Sequence 536 AA;
XX
XX Query Match          97.0%; Score 3014; DB 3; Length 536;
XX Best Local Similarity 100.0%; Pred. No. 1e-210;
XX Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 SIKRDNDYKSNPMRIVCYVGTWSVYHKVDPTYTIEDIDPFKCTHLMYGFADIKYKTIQ 79
DB 1 SIKRDNDYKSNPMRIVCYVGTWSVYHKVDPTYTIEDIDPFKCTHLMYGFADIKYKTIQ 60
QY 80 VFDPYQDDNHNSEKRGYERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANPTVROOF 139
DB 61 VFDPYQDDNHNSEKRGYERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANPTVROOF 120
QY 140 IQSVLDFLOEYKFDGLDWEYFGSLGNPKDKQNYLALVRELKDAFPHGYLLTAAYS 199
DB 121 IQSVLDFLOEYKFDGLDWEYFGSLGNPKDKQNYLALVRELKDAFPHGYLLTAAYS 180
QY 200 PGKDKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNVN 259
DB 181 PGKDKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNVN 240
QY 260 YTHYIYLNNGATRDKLVMGVFFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEEGVLSTI 319
DB 241 YTHYIYLNNGATRDKLVMGVFFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEEGVLSTI 300
QY 320 ELQOLFQKEBWHIYDEYNNAPYGYNDKIWGVYDDLASISCKLAFELKELGVSGVMVWSLE 379
DB 301 ELQOLFQKEBWHIYDEYNNAPYGYNDKIWGVYDDLASISCKLAFELKELGVSGVMVWSLE 360
QY 380 NDDFKGHCGRKPNLLKNVHNKNGDEKNSPECILGPSTTPTPTPTPTPTPTPTPTPTPTPT 439
DB 361 NDDFKGHCGRKPNLLKNVHNKNGDEKNSPECILGPSTTPTPTPTPTPTPTPTPTPTPTPT 420
QY 440 TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 499
DB 421 TPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 480
QY 500 VDGHLLKCYKEGDIHPHTNTHKYLVECFVNGGWVHIMCPPTGIIWCKLTCIGE 555
DB 481 VDGHLLKCYKEGDIHPHTNTHKYLVECFVNGGWVHIMCPPTGIIWCKLTCIGE 536
XX
XX RESULT 5
XX AAU96329
XX ID AAU96329 standard; protein; 536 AA.

```

```

XX AC AAU96329;
XX DT 15-JUL-2002 (first entry)
XX DE Der HWM-map polypeptide #16.
XX KW Der HWM-map; American house dust mite; antiallergic; mite; IgE;
XX KW mite allergenic protein; immunoglobulin E; hypersensitivity;
XX OS immunocomplex formation.
XX OS Dermatophagoides farinae.
XX PN WO200222807-A2.
XX PD 21-MAR-2002.
XX PF 14-SEP-2001; 2001WO-US028730.
XX PR 14-SEP-2000; 2000US-00662293.
XX PA (HESK-) HESKA CORP.
XX PI McCall CA, Hunter SW, Weber ER;
XX DR WPI; 2002-351888/38.
XX DR N-PSDB; ABK69575.
XX PT New mite allergenic protein isolated from Dermatophagoides, designated
XX PT Der HWM-map protein, useful as a vaccine for treating mite allergy.
XX PS Claim 12; Page 125-127; 161pp; English.
XX CC The invention relates to an isolated mite allergenic protein of
XX CC Dermatophagoides, designated Der HWM-map protein, and its related nucleic
XX CC acid. The Der HWM-map protein is useful for eliciting an immune response
XX CC against Der HWM-map protein. The protein or a reagent comprising a non-
XX CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
XX CC cat) susceptible to or having an allergic response to a mite. A
XX CC therapeutic composition is useful for desensitising a host animal to an
XX CC allergic response to a mite. The DNA and protein can be used in the
XX CC detection of anti-Der HWM-map antibodies in animal fluids, and inhibition
XX CC of immunoglobulin (Ig)E or Der HWM-map protein activity associated with a
XX CC disease. Antibodies that bind to Der HWM-map are useful for inhibiting
XX CC binding of proteins to IgE, to prevent immunocomplex formation, thus
XX CC reducing hypersensitivity responses to mite allergens, and as vaccines
XX CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
XX CC represent Der HWM-map polypeptides of the invention
XX SQ Sequence 536 AA;
XX
XX Query Match          97.0%; Score 3014; DB 5; Length 536;
XX Best Local Similarity 100.0%; Pred. No. 1e-210;
XX Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 20 SIKRDNDYKSNPMRIVCYVGTWSVYHKVDPTYTIEDIDPFKCTHLMYGFADIKYKTIQ 79
DB 1 SIKRDNDYKSNPMRIVCYVGTWSVYHKVDPTYTIEDIDPFKCTHLMYGFADIKYKTIQ 60
QY 80 VFDPYQDDNHNSEKRGYERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANPTVROOF 139
DB 61 VFDPYQDDNHNSEKRGYERFNNRLKNPELTTMISLGGWYEGSEKYSDMAANPTVROOF 120
QY 140 IQSVLDFLOEYKFDGLDWEYFGSLGNPKDKQNYLALVRELKDAFPHGYLLTAAYS 199
DB 121 IQSVLDFLOEYKFDGLDWEYFGSLGNPKDKQNYLALVRELKDAFPHGYLLTAAYS 180
QY 200 PGKDKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNVN 259
DB 181 PGKDKIDRAYDIKELNKLFDWMNVMTYDYGHWENFYGHNAPLYKRPDETDELHTYFNVN 240
QY 260 YTHYIYLNNGATRDKLVMGVFFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEEGVLSTI 319

```

Db 241 YTMHYLNGATRDKLWGVPPYGRAMSIEDRSKULGDPAGKMPGFIISGEGVLSYI 300
Qy 320 ELCCQLFQKEEWHIQDEYNNAPYNDKIWGYDDLASISCKLAFKLKELGSGVWVMSLE 379
Db 301 ELCCQLFQKEEWHIQDEYNNAPYNDKIWGYDDLASISCKLAFKLKELGSGVWVMSLE 360
Qy 380 NDDFKGHCQKPNLKNVNMINGDEKNSFECILGSPSTTPPTPTPTPTPTPTPTPTPTPT 439
Db 361 NDDFKGHCQKPNLKNVNMINGDEKNSFECILGSPSTTPPTPTPTPTPTPTPTPTPTPT 420
Qy 440 TPTTPT 499
Db 421 TPTTPT 480
Qy 500 VDGHLIKCYKEGDIHPHTNIHKYLVCVFVNGGWVHIMPCPGTIWCQEKLTICGE 555
Db 481 VDGHLIKCYKEGDIHPHTNIHKYLVCVFVNGGWVHIMPCPGTIWCQEKLTICGE 536

RESULT 6
AAV52533
ID AAV52533 standard; protein; 509 AA.
AC
XX AAV52533;
XX
DT 06-AUG-2003 (revised)
DT 22-FEB-2000 (first entry)
XX
D. pteronyssius 98 kD mite allergen protein (map) PDerp98-509.
XX
XX Mite allergen protein; map; high molecular weight; HMW-map; allergy;
KW house dust mite; IgE; immunoglobulin E; allergen; mapB;
KW hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
KW canine; veterinary; antibody; vaccine; immunisation.
XX
OS Dermatophagoides pteronyssinus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19 /note= "Signal peptide"
FT FT
FT Protein 20..509 /note= "Mature PDerp98-509"
XX
PN WO9954349-A2.
XX
XX 28-OCT-1999.
XX
XX 16-APR-1999; 99WO-US008524.
XX
XX 17-APR-1998; 98US-00062013.
XX 13-MAY-1998; 98US-0085295P.
XX 02-SEP-1998; 98US-0098909P.
XX
XX (HESK-) HESKA CORP.
XX
XX Mccall CA, Hunter SW, Weber ER;
XX
XX WPI; 2000-052700/04.
XX N-PSDB; AAZ38585, AAZ38586, AAZ38587, AAZ38588.
XX
XX Novel high molecular weight Dermatophagoides nucleic acid polypeptides
XX used to modify an animals' hypersensitivity to mite allergens.
XX
XX Claim 3; Page 134-136; 154pp; English.
XX

This sequence represents Dermatophagoides pteronyssius mite allergen protein (map) PDerp98-509. PDerp98-509 has a molecular weight of 98 kD, comprising 509 amino acids, and has a high degree of homology with the D. farinae 98 kD allergen, mapB (AAV52523). Nucleic acid molecules encoding PDerp98-509 were isolated from a D. pteronyssius cDNA library by hybridisation with a probe encoding the D. farinae high molecular weight map (HMW-map) composition. Mite allergenic proteins and peptides, and nucleic acids encoding them, may be used in therapeutic compositions to

CC modify an animal's hypersensitivity reaction to mite allergens. Animals
CC that may be treated include mammals and birds, especially felines,
CC canines, equines, humans, other pets, and work or domestic animals. The
CC proteins or fragments may also be used to diagnose allergies via a skin
CC test. The proteins and peptides can also be used to raise antibodies,
CC which have a variety of potential uses. For example, they can be used as
CC vaccines to passively immunise animals against dust mite
CC hypersensitivity, as positive controls in test kits and as tools to
CC recover desired dust mite allergens from a mixture of proteins. (Updated
CC on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 509 AA;

Query Match 81.8%; Score 2542; DB 3; Length 509;
Best Local Similarity 81.7%; Pred. No. 2e-176;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;
Qy 1 MKTTFALFCIWACIGLNNATKRHHNYSKNPWRIVCVGTWSVYHKVDPYTIEDIDPFK 60
Db 1 MKTTFALFCIWACIGLNNATKRHHNYSKNPWRIVCVGTWSVYHKVDPYTIEDIDPFK 60
Qy 61 CTHLMYGFPAKIDYKTIQVDPYQDDNHNHSEKRGYERFNNLRLKNPELTMTISLGWY 120
Db 61 CTHLMYGFPAKIDYKTIQVDPYQDDNHNHSEKRGYERFNNLRLKNPELTMTISLGWY 120
Qy 121 EGSEKYSDMAANPTYROQFIQSVDLFLQYKFDGLDWEYPGSRLGNPKIDKQNYLTV 180
Db 121 EGSEKYSDMAANPTYROQFIQSVDLFLQYKFDGLDWEYPGSRLGNPKIDKQNYLTV 180
Qy 181 RELKDAFPHGYLITAAVSPGDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
Db 181 RELKDAFPHGYLITAAVSPGDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
Qy 241 PLYKPEPDTDELHYENVNNTMYHLLANGATRDKLWGVPPYGRAWSIEDRSKVLGDP 300
Db 241 PLYKPEPDTDELHYENVNNTMYHLLANGATRDKLWGVPPYGRAWSIEDRSKVLGDP 300
Qy 301 KGMSPPGTISGEGVLSYIELCOLFQKEEWHIQDEYNNAPYNDKIWGYDDLASISC 360
Db 301 KGMSPPGTISGEGVLSYIELCOLFQKEEWHIQDEYNNAPYNDKIWGYDDLASISC 360
Qy 361 KLAFKLKELGSGVWVMSLENDDFKHCQKPNLKNVNMINGDEKNSFECILGSPSTTP 420
Db 361 KLAFKLKELGSGVWVMSLENDDFKHCQKPNLKNVNMINGDEKNSFECILGSPSTTP 420
Qy 421 TPTTPT 480
Db 421 TPTTPT 434
Qy 481 TPTTPT 537
Db 435 ---PTTDTSTSETPKYTIYDGHLLIKCYKQGYLPHPTDVHKYLVCYIATFNGGWVHIM 491

RESULT 7
AAU96337
ID AAU96337 standard; protein; 509 AA.
XX
XX AAU96337;
XX
XX 15-JUL-2002 (first entry)
XX
XX Der HMW-map polypeptide #24.
XX
XX Der HMW-map; American house dust mite; antiallergic; mite; IgE;
KW mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
XX Dermatophagoides farinae.
OS

XX WO200222807-A2.
XX 21-MAR-2002.
XX 14-SEP-2001; 2001WO-US028730.
XX 14-SEP-2000; 2000US-00662293.
XX (HESK-) HESKA CORP.
XX McCall CA, Hunter SW, Weber ER;
XX WPI; 2002-351888/38.
XX N-PSDB; ABK69581.
XX New mite allergenic protein isolated from Dermatophagoides, designated
XX Der HMW-map protein, useful as a vaccine for treating mite allergy.
XX Claim 12; Page 134-136; 161pp; English.
XX The invention relates to an isolated mite allergenic protein of
XX Dermatophagoides, designated Der HMW-map protein, and its related nucleic
XX acid. The Der HMW-map protein is useful for eliciting an immune response
XX against Der HMW-map protein. The protein or a reagent comprising a non-
XX proteinaceous epitope is useful for identifying an animal (e.g., dog,
XX cat) susceptible to or having an allergic response to a mite. A
XX therapeutic composition is useful for desensitising a host animal to an
XX allergic response to a mite. The DNA and protein can be used in the
XX detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
XX of immunoglobulin (IgE) or Der HMW-map protein activity associated with a
XX disease. Antibodies that bind to Der HMW-map are useful for inhibiting
XX binding of proteins to IgE, to prevent immunocomplex formation, thus
XX reducing hypersensitivity responses to mite allergens, and as vaccines
XX against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
XX represent Der HMW-map polypeptides of the invention
XX
SQ Sequence 509 AA;
Query Match 81.8%; Score 2542; DB 5; Length 509;
Best Local Similarity 81.7%; Pred. No. 2e-176;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;
QY 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWSVYHKVDPTIEDIDPFK 60
DB 1 MKTIFALFCIWACIGLMNAATKRDHNNYSKNPMRIVCVGTWSVYHKVDPTIEDIDPFK 60
QY 61 CTHLMYGFAXIDEYKTIQVDFPDQDDHNSWEKRGYERFNNLRKNPELTMTISLGGWY 120
DB 61 CTHLMYGFAXIDEYKTIQVDFPDQDDHNSWEKRGYERFNNLRKNPELTMTISLGGWY 120
QY 121 EGSEKYSMDMANPTYRQFQVSLDFLOEYKFDGLDWEYPSRGLGNPKDKQNYLTV 180
DB 121 EGSEKYSMDMANPTYRQFQVSLDFLOEYKFDGLDWEYPSRGLGNPKDKQNYLTV 180
QY 181 RELKDAEPFHGYLLTAAVSPGKIDRAYDIKELNKLFDWMNVMTYDYGWENFVGHNA 240
DB 181 RELKDAEPFHGYLLTAAVSPGKIDRAYDIKELNKLFDWMNVMTYDYGWENFVGHNA 240
QY 241 PLYKRPDETDLHYTFVNYMTMYLLNNGATRDKLVMGVFPFYGRAWSIEDRSKILGDDPA 300
DB 241 PLYKRPDETDLHYTFVNYMTMYLLNNGATRDKLVMGVFPFYGRAWSIEDRSKILGDDPA 300
QY 301 KGMSPPGFISGEBVLSYIELCOLFOKEWHIQDEYNNAPYGNDKIWWGYDDLASISC 360
DB 301 KGMSPPGFITGEGVLSYIELCOLFOKEWHIQDEYNNAPYGNDKIWWGYDDLASISC 360
QY 361 KLAFKELGVSGVMVWSLENDDEKHCCKPNPLKNVHNMINGDEKNSFECILGPSSTTP 420
DB 361 KLAFKELGVSGVMVWSLENDDEKHCCKPNPLKNVHNMINGDEKNSFECILGPSSTTP 420
QY 421 TPTTPT 480
DB 421 TPTTPT 480

DB 421 TPTTPT 434
QY 481 TPTTPT 537
DB 435 ---PTTDDSTSTPKYTYIDGHLIKYKOGYLPDPTDVHKYLVCEVIATPNGGWWHIM 491
QY 538 PCPPGTIWCQEKLTICGE 555
DB 492 DCPKGTWHTLKNKICQE 509
RESULT 8
AAU96338
ID AAU96338 standard; protein; 509 AA.
XX AC AAU96338;
XX 15-JUL-2002 (first entry)
XX Der HMW-map polypeptide #25.
XX Der HMW-map; American house dust mite; antiallergic; mite; IgE;
XX mite allergenic protein; immunoglobulin E; hypersensitivity;
XX immunocomplex formation.
XX Dermatophagoides farinae.
XX WO200222807-A2.
XX 21-MAR-2002.
XX 14-SEP-2001; 2001WO-US028730.
XX 14-SEP-2000; 2000US-00662293.
XX (HESK-) HESKA CORP.
XX McCall CA, Hunter SW, Weber ER;
XX WPI; 2002-351888/38.
XX N-PSDB; ABK69583.
XX New mite allergenic protein isolated from Dermatophagoides, designated
XX Der HMW-map protein, useful as a vaccine for treating mite allergy.
XX Claim 12; Page 139-141; 161pp; English.
XX The invention relates to an isolated mite allergenic protein of
XX Dermatophagoides, designated Der HMW-map protein, and its related nucleic
XX acid. The Der HMW-map protein is useful for eliciting an immune response
XX against Der HMW-map protein. The protein or a reagent comprising a non-
XX proteinaceous epitope is useful for identifying an animal (e.g., dog,
XX cat) susceptible to or having an allergic response to a mite. A
XX therapeutic composition is useful for desensitising a host animal to an
XX allergic response to a mite. The DNA and protein can be used in the
XX detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
XX of immunoglobulin (IgE) or Der HMW-map protein activity associated with a
XX disease. Antibodies that bind to Der HMW-map are useful for inhibiting
XX binding of proteins to IgE, to prevent immunocomplex formation, thus
XX reducing hypersensitivity responses to mite allergens, and as vaccines
XX against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
XX represent Der HMW-map polypeptides of the invention
XX
SQ Sequence 509 AA;
Query Match 81.8%; Score 2542; DB 5; Length 509;
Best Local Similarity 81.7%; Pred. No. 2e-176;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;
QY 1 MKTIVAILSIMACIGLMNASIKRDNDYSKNPMRIVCVGTWSVYHKVDPTIEDIDPFK 60
DB 1 MKTIFALFCIWACIGLMNAATKRDHNNYSKNPMRIVCVGTWSVYHKVDPTIEDIDPFK 60

QY 61 CTHLMGFAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWY 120
DB 61 CTHLMGFAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWY 120
QY 121 EGSEKSDMAANTYRQOFTQSVLDFLQYKFGDLWDWEYFPGSRGNPKIDKQNTYALV 180
DB 121 EGSEKSDMAANTYRQOFTQSVLDFLQYKFGDLWDWEYFPGSRGNPKIDKQNTYALV 180
QY 181 RELKDAPEPHGYLLTAASVPGKDKIDRAYDIKELNKLFDMMNVTYDYGHWENFYGHNA 240
DB 181 RELKDAPEPHGYLLTAASVPGKDKIDRAYDIKELNKLFDMMNVTYDYGHWENFYGHNA 240
QY 241 PLYKRPDETDELHTYFNVTMTHYLLANGATRDKLVMGVFFYGRAWSIEDRSKLGDPDA 300
DB 241 PLYKRPDETDELHTYFNVTMTHYLLANGATRDKLVMGVFFYGRAWSIEDRSKLGDPDA 300
QY 301 KGMSPGFGISGEGVLSYIELCOLFOKEWHIOYDEYNAPYGNKIDKQNTYALV 360
DB 301 KGMSPGFGISGEGVLSYIELCOLFOKEWHIOYDEYNAPYGNKIDKQNTYALV 360
QY 361 KLAFLKELGSGVMVWSLENDPKHGCGPNPLNKNVHNMINGDEKNSYECILGPSTTTP 420
DB 361 KLAFLKELGSGVMVWSLENDPKHGCGPNPLNKNVHNMINGDEKNSYECILGPSTTTP 420
QY 421 TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
DB 421 TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
QY 481 TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 537
DB 435 ---PTTDTSTSTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 537
QY 538 PCPPGTIWCOKTCTIGE 555
DB 492 DCPKGRWHTATLNCIQE 509

RESULT 9
AA52535
ID RAY52535 standard; protein; 490 AA.
XX
AC AA52535;
XX
XX
DT 06-AUG-2003 (revised)
DT 22-FEB-2000 (first entry)
XX
DE D. pteronyssius 98 kD mite allergen protein (map) Pderp98-490.
XX
XX Mite allergen protein; map; high molecular weight; HMW-map; allergy;
XX house dust mite; IgE; immunoglobulin E; allergen; mapB;
XX hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
XX canine; veterinary; antibody; vaccine; immunisation.
XX
OS Dermatophagoides pteronyssinus.
XX
FH Key Location/Qualifiers
FT Modified-site 115..117
FT /note= "Asn is N-glycosylated"
FT Modified-site 240..242
FT /note= "Asn is N-glycosylated"
XX
PN WO9954349-A2.
XX
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-US008524.
XX
PR 17-APR-1998; 98US-00062013.
PR 13-MAY-1998; 98US-0085295P.
PR 02-SEP-1998; 98US-0098909P.
XX
XX (HESK-) HESKA CORP.
XX
XX

PI McCall CA, Hunter SW, Weber ER;
XX
DR WPI: 2000-052700/04.
DR N-PSDB; AAZ38589, AAZ38590.
XX
XX Novel high molecular weight Dermatophagoides nucleic acid polypeptides
PT used to modify an animals' hypersensitivity to mite allergens.
XX
XX Claim 3; Page 147-149; 154pp; English.
XX
XX This sequence represents Dermatophagoides pteronyssius mite allergen
CC protein (map) Pderp98-490, the mature form of pDerp98-509. Pderp98-490
CC has a molecular weight of 98 kD, comprising 490 amino acids, and has a
CC high degree of homology with the D. farinae mature 98 kD allergen, mapB
CC (AAV52525). Nucleic acid molecules encoding Pderp98-490 were isolated
CC from a D. pteronyssius cDNA library by hybridisation with a probe
CC encoding the D. farinae high molecular weight map (HMW-map) composition.
CC Mite allergenic proteins and peptides, and nucleic acids encoding them,
CC may be used in therapeutic compositions to modify an animal's
CC hypersensitivity reaction to mite allergens. Animals that may be treated
CC include mammals and birds, especially felines, canines, equines, humans,
CC other pets, and work or domestic animals. The proteins or fragments may
CC also be used to diagnose allergies via a skin test. The proteins and
CC peptides can also be used to raise antibodies, which have a variety of
CC potential uses. For example, they can be used as vaccines to passively
CC immunise animals against dust mite hypersensitivity, as positive controls
CC in test kits and as tools to recover desired dust mite allergens from a
CC mixture of proteins. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 490 AA;
Query Match 79.7%; Score 2475; DB 3; Length 490;
Best Local Similarity 82.5%; Pred. No. 1.5e-171;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
QY 22 KRDNHNSKPMRIVCVYVGTWVYKVDPTIEDIDPKCTHLMYGAIDYKTYIQVF 81
DB 3 KRDNHNSKPMRIVCVYVGTWVYKVDPTIEDIDPKCTHLMYGAIDYKTYIQVF 62
QY 82 DPYQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWYEGSEKYSMAANTYRQOFTQ 141
DB 63 DPFQDDNHNSEKRGYERFNNLRKNPELTMTISLGGWYEGSEKYSMAANTYRQOFTQ 122
QY 142 SVLDFLQYKFGDLWDWEYFPGSRGNPKIDKQNTYALVRELKDAPEPHGYLLTAASVPG 201
DB 123 SVLDFLQYKFGDLWDWEYFPGSRGNPKIDKQNTYALVRELKDAPEPHGYLLTAASVPG 182
QY 202 KDKIDRAYDIKELNKLFDMMNVTYDYGHWENFYGNAPLYKRPDETDELHTYFNVT 261
DB 183 KDKIDRAYDIKELNKLFDMMNVTYDYGHWENFYGNAPLYKRPDETDELHTYFNVT 242
QY 262 MHYLLANGATRDKLVMGVFFYGRAWSIEDRSKLGDPKGMSPGFGISGEGVLSYIEL 321
DB 243 MHYLLANGATRDKLVMGVFFYGRAWSIEDRSKLGDPKGMSPGFGISGEGVLSYIEL 302
QY 322 COLFOKEWHIOYDEYNAPYGNKIDKQNTYALVRELKDAPEPHGYLLTAASVPG 381
DB 303 COLFOKEWHIOYDEYNAPYGNKIDKQNTYALVRELKDAPEPHGYLLTAASVPG 362
QY 382 DFKHGCGPNPLNKNVHNMINGDEKNSYECILGPSTTPTTPTTPTTPTTPTTPTTPTT 441
DB 363 DFKHGCGPNPLNKNVHNMINGDEKNSYECILGPSTTPTTPTTPTTPTTPTTPTTPTT 415
QY 442 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 501
DB 416 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 433
QY 502 GHLIKCYKEGDI PHPTNHLKYLVCDFV---NGGWWHIMPCPGTINCOEKLTCIGE 555
DB 434 GHLIKCYKEGDI PHPTNHLKYLVCDFV---NGGWWHIMPCPGTINCOEKLTCIGE 490

AAU96339
ID AAU96339 standard; protein; 490 AA.
XX
AC AAU96339;
XX
DT 15-JUL-2002 (first entry)
XX
DE Der HMW-map polypeptide #26.
XX
KW Der HMW-map; American house dust mite; antiallergic; mite; IgE;
XX mite allergenic protein; immunoglobulin E; hypersensitivity;
KW immunocomplex formation.
XX
OS Dermatophagoides farinae.
XX
PN WO200222807-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028730.
XX
PR 14-SEP-2000; 2000US-00662293.
XX
PA (HESK-) HESKA CORP.
XX
PI McCall CA, Hunter SW, Weber ER;
XX
DR WPI; 2002-351888/38.
DR N-PSDB; ABK69585.
XX
PT New mite allergenic protein isolated from Dermatophagoides, designated
PT Der HMW-map protein, useful as a vaccine for treating mite allergy.
XX
PS Claim 12; Page 144-146; 161pp; English.
XX
CC The invention relates to an isolated mite allergenic protein of
CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
CC acid. The Der HMW-map protein is useful for eliciting an immune response
CC against Der HMW-map protein. The protein or a reagent comprising a non-
CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
CC cat) susceptible to or having an allergic response to a mite. A
CC therapeutic composition is useful for desensitising a host animal to an
CC allergic response to a mite. The DNA and protein can be used in the
CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
CC binding of proteins to IgE, to prevent immunocomplex formation, thus
CC reducing hypersensitivity responses to mite allergens, and as vaccines
CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
CC represent Der HMW-map polypeptides of the invention
XX
SQ Sequence 490 AA;

Query Match 79.7%; Score 2475; DB 5; Length 490;
Best Local Similarity 82.5%; Pred. No. 1.5e-171;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;

QY 22 KRDNDYKSNPMRIVCYGTVSWVHKVDPTIEDIDPKCTHLMYGFADKDEYKTIQVF 81
DB 3 KRDENYKSNPMRIVCYGTVSWVHKVDPTIEDIDPKCTHLMYGFADKDEYKTIQVF 62

QY 82 DPYODNHSNEKRGYERFNNLRKLNPELTMTWISLGGWYEGSEKYSMDAANPTVROQFIQ 141
DB 63 DFDNDSNHSNEKRGYERFNNLRKLNPELTMTWISLGGWYEGSEKYSMDAANPTVROQFVQ 122

QY 142 SVLDFLOQYKFDGLDWEYFGSRNLGNPKIDKQNYLALVRELKDAFEPHYLLTAASVPG 201
DB 123 SVLDFLOQYKFDGLDWEYFGSRNLGNPKIDKQNYLALVRELKDAFEPHYLLTAASVPG 182

QY 202 KDKIDRAVDIKELNKLFWNNVWYDYGWENFYGHNAPIYKRPDETDELHTFNNVNYT 261
DB 183 KDKIDVAVELKELNQLFDWNNVWYDYGWENFYGHNAPIYKRPDETDELHTFNNVNYT 242

QY 262 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLKLGDPAGKMGPPGFISGEGVLSYIEL 321
DB 243 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLKLGDPAGKMGPPGFITGEEGVLSYIEL 302

QY 322 COLFQKEEWHIOYDEYVYNDKLVGVDDIASISCKLAFKELGVSGVWVWSLEND 381
DB 303 COLFQKEEWHIOYDEYVYNDKLVGVDDIASISCKLAFKELGVSGVWVWSLEND 362

QY 382 DFKGHCQKPNLLNKNVHNMINGDEKNSFECILGSPSTTTPTPTPTPTPTPTPTPTPTPT 441
DB 363 DFKGHCQKPNLLNKNVHNMINGDEKNSVECLGSPSTTTPTPTPTPTPTPTPTPTPTPT 415

QY 442 TTTPSPPT 501
DB 416 -----PTTDTSTSTPKYTYID 433

QY 502 GHLIKCYKEGDIPTHTNIHKYLVCFFV---NGGWHVHIMPCPPGTINCOEKLTCIGE 555
DB 434 GHLIKCYKQGYLPHTDVKHYLVCEYIATPNGGWHVHIMDCPKGRWHATLKNLCIQE 490

RESULT 11
ABP72636
ID ABP72636 standard; protein; 525 AA.
XX
AC ABP72636;
XX
DT 11-JUN-2003 (first entry)
XX
DE Anopheles gambiae chitinase.
XX
KW Insecticide; pesticide; insect control; insect; toxin; chitinase; enzyme;
KW neuropeptide; transgenic plant; crop protection; mosquito.
XX
OS Anopheles gambiae.

FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /label= Signal_peptide
FT /label= Mature_protein
FT Misc-difference 31
FT /note= "possible trypsin activation site"
FT Misc-difference 32
FT /note= "possible trypsin activation site"
FT Active-site 150..157
FT Domain 401..466
FT /note= "serine/threonine/proline-rich domain"
XX WO2003014150-A2.
XX
PD 20-FEB-2003.
XX
PF 06-AUG-2002; 2002WO-GB003598.
XX
PR 08-AUG-2001; 2001GB-00019274.
XX
PA (UYDU-) UNIV DURHAM.
PA (ENV1-) DEPT ENVIRONMENT FOOD & RURAL AFFAIRS.
XX
PI Gatehouse JA, Fitches EC, Edwards JP;
XX WPI; 2003-278469/27.
DR N-PSDB; ABZ81875.
XX
PT Fusion protein useful for combating insect pests, comprises a
PT translocating moiety comprising a plant protein capable of acting as a
PT carrier to translocate toxic moiety inside plant pathogen, and a toxic
PT moiety.
XX
PS Claim 7; Fig 15; 51pp; English.
XX
CC The present sequence is that of the mosquito Anopheles gambiae chitinase.

CC This protein is used in claimed fusion proteins of the invention. Such
 CC fusion proteins comprise a translocating moiety and a toxic moiety, where
 CC the translocating moiety is a plant protein (e.g. a lectin) capable of
 CC acting as a carrier to translocate the toxic moiety across the gut wall
 CC of a plant pathogen, and the toxic moiety is an insect-derived peptide or
 CC protein capable of causing deleterious effects on growth, development,
 CC reproduction or mortality in pest insects. Suitable insect peptides and
 CC proteins include allatotaxins, chitinase, diuretic hormone and their
 CC metabolites and analogues. Polynucleotides encoding the fusion protein,
 CC vectors, host cells and transgenic plants that are resistant to disease
 CC are also provided. The fusion protein is target-specific, and resists
 CC degradation in the insect gut

SQ Sequence 525 AA;

Query Match 32.5%; Score 1008.5; DB 6; Length 525;
 Best Local Similarity 38.4%; Pred. No. 7.7e-65;
 Matches 211; Conservative 93; Mismatches 187; Indels 59; Gaps 13;

QY 8 LSIMACIGLMMASIKRD-HNDYSKNPMRLVCVGTWMSYHKVD-PYTIEDDPKCTHLM 65
 DB 5 VGVLLVLAFAAEAEHPKAASAEKKVVCVGTWMAVYRFGNGRYDIEH:DPSCLTML 64
 QY 66 YGFAIDEXYTIQVDPYQDDNHNHSEKRGYERFNNRLKNPELTTMISLGGWYEGSEK 125
 DB 65 YGFFGINE-DAVRIIDPDLLEN-WGRGHKRVGLXNVGPGCLTLAAGGWNESGRK 122
 QY 126 YSDMAANPTTROOFTQSVLDLFQEQYKFDGLDWEYPGSRGNPKIDKONYALVRELKD 185
 DB 123 PSMAASGELRKRFSIDCVAFQCRHGFDGIDLDWEYPAQRDGNPLIDRDNAQLVEEMRE 182
 QY 186 APEPHGYLLTAAVSPGKKIDRAYDIKELNKLFDWNNVTYDGGWENFYGHNAPLYKR 245
 DB 183 EFDHYGLLLTAAVSVEFSAGVSYDIPRISKSFHFLNVMVYDMHGAWSDYCGINAPLYRG 242
 QY 246 PDETDELHYTFNNVTMYLNNNGATROKLVGVFPFYGRANSIEDRSKLGDDPAKGNP 305
 DB 243 SADTTDLRGQIINVASIHFMLAQGGTGRKLVGLIPLYGRNFTLASAANTQIGAPTGGGT 302
 QY 306 PGFISGEEGLSVIELCOLFOKEWHIOYDEVYNAPYGVNDKIWVGVDLGLASISCKLAF 365
 DB 303 VGRYTRPGVWGYNEFECEKLAETANDLWSEQQVYAVRNNQWVGYYDDLSVQLKVXL 362
 QY 366 KELGVSGVMVWLENDDPKGHC-GPKNPLLNKHVNMINGDEKNSPECILGPSTPTPTPT 424
 DB 363 LDQGLGAMVMSLETDDPLGVCGGGRYPLMHEIRSLVNGGT-----PSTTTPPSV 413
 QY 425 TPT 484
 DB 414 APTT-----STVAPGTTTPTTGANPGTTPPT--SDAPNHTTSTTTTEGNPGTTRPPSG 466
 QY 485 TTEHTSETPKYTYVYDGLIKCY--KEGDIHPNTNHHKLYC-----EFVNGGWVH 535
 DB 467 -----DG---PCAGRGYGVFPHPTNCARYICTADTYEFT----- 500
 QY 536 IMPCPPTGIW 545
 DB 501 ---CPPTGLP 507

RESULT 12

ABB58595
 ID ABB58595 standard; protein; 4498 AA.
 XX
 AC ABB58595;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2577.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL02698.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

XX Disclosure; SEQ ID NO 2577; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 4498 AA;

Query Match 27.5%; Score 854.5; DB 4; Length 4498;
 Best Local Similarity 33.8%; Pred. No. 2e-52;
 Matches 202; Conservative 88; Mismatches 188; Indels 119; Gaps 21;

QY 2 KTIYAILISACIGLMMASIKRDHNDYSKNPMRLVCVGTWMSYH---KVDYPTIEDID 57
 DB 33 QTLFLCALAYCI-----NEASSEG-RVVCYTTWMSYRPTAKFNP---QNIN 77
 QY 58 PFKCTHLMY---GPAKIDEXYTIQVDPYQDDNHNHSEKRGYERFNNRLKNPELTTMI 114
 DB 78 PYLCTHUVYAFGGFTKDNQMK---PDKYQD-----IEGGYAKFTGLKTYNKQLXTMI 128
 QY 115 SLGWYEGSBKYSDMAANPTTROOFTQSVLDLFQEQYKFDGLDWEYPGSRGLNPKIDKQ 174
 DB 129 AIGGNWASRSPPLVASNERRQOFIKNLIKFLRQNHFDGIDLDWEYPAHREGSKSRD 188
 QY 175 NYLALVRELKDAFEPHG-----YLTAAVSPGKKIDRAYDIKELNKLFDWNNVTYD 227
 DB 189 NYAQVQELRAEFERAEKTRTLLTMAVPAGIEYIDKGYDVPKLYKLDWNFNLTYD 248
 QY 228 YHGWENFYGHNAPLYKRDPDETDELHYT---FNNVTMYLNNNGATROKLVGVFPYGR 284
 DB 249 FHSSEPSVNHAPLYSL--EEDSEYNYDAELNIDYSIKYLLKAGADRDKLVLGIFTYGR 306
 QY 285 AWSIEDRSKLGIDPAKGNMPPPGFISGEEGLSVIELCOLFOKE-EWH-IQYDEVYNAPY 342
 DB 307 SYTLNEESTELGAPAGPGEGQGDATREKGYLAYEICQTLKDDPEVTVPQNVNMGPY 366
 QY 343 GYNDKIWVGVDLGLASISCKLAFKELGVSGVMVWLENDDPKGHC-GPKNPLLNKHVNM 401
 DB 367 AYRENQWVGVDDEAIVRKAEEYVVAQGLGGMFWAIDNDDFRGTCNGKPYLIEAA---- 422
 QY 402 NGDEKNSFECILGPSTPT 445
 DB 423 ----KEAMVEALGLGINEVAKPSGPKPSRSRSDNASNRNLNGKTEAPLSSRRPSATR 478

| | | |
|----|---|--------------------------------|
| XX | Manduca sexta chitinase. | |
| DE | | |
| XX | | |
| XX | Insecticide; peptidase; insect control; insect; toxin; chitinase; enzyme; | |
| KW | neuropeptide; transgenic plant; crop protection. | |
| XX | | |
| XX | | |
| OS | Manduca sexta. | |
| XX | | |
| XX | Location/Qualifiers | |
| FT | Key | |
| FT | Peptide | 1. .19 |
| FT | | /label= Signal_peptide |
| FT | Protein | 20. .554 |
| FT | | /label= Mature_protein |
| FT | | 85. .88 |
| FT | Modified-site | /note= "Asn is N-glycosylated" |
| FT | | 138. .146 |
| FT | Active-site | |
| FT | | 303. .306 |
| FT | Modified-site | /note= "Asn is N-glycosylated" |
| FT | | 545. .548 |
| FT | Modified-site | /note= "Asn is N-glycosylated" |
| FT | | |

| | |
|----|---|
| AA | WO2003014150-A2. |
| PN | |
| XX | XX |
| XX | XX |
| XX | 20-FEB-2003. |
| XX | XX |
| XX | XX |
| XX | 06-AUG-2002; 2002WO-GB003598. |
| XX | XX |
| XX | XX |
| XX | 08-AUG-2001; 2001GB-00019274. |
| XX | XX |
| XX | (UYDU-) UNIV DURHAM. |
| PA | (ENVI-) DEPT ENVIRONMENT FOOD & RURAL AFFAIRS. |
| PA | |
| XX | XX |
| XX | Gatehouse JA, Fitches EC, Edwards JP; |
| XX | |
| XX | WPI; 2003-278469/27. |
| DR | XX |
| DR | XX |
| XX | XX |
| PT | Fusion protein useful for combating insect pests, comprises a |
| PT | PT translocating moiety comprising a plant protein capable of acting as a |
| PT | carrier to translocate toxic moiety inside plant pathogen, and a toxic |
| PT | moistv. |

| | | |
|----|---|--|
| | Claim 7; Fig 10; Sipp; English. | |
| XX | The present sequence is that of Manduca sexta chitinase. This protein can | |
| PS | be used in claimed fusion proteins or the invention. Such fusion proteins | |
| XX | comprise a translocating moiety and a toxic moiety, where the | |
| CC | translocating moiety is a plant protein (e.g. a lectin) capable of acting | |
| CC | as a carrier to translocate the toxic moiety across the gut wall of a | |
| CC | plant pathogen, and the toxic moiety is an insect-derived peptide or | |
| CC | protein capable of causing deleterious effects on growth, development, | |
| CC | reproduction or mortality in pest insects. Suitable insect peptides and | |
| CC | proteins include allatostatin, chitinase, diuretic hormone and their | |
| CC | metabolites and analogues. Polynucleotides encoding the fusion protein, | |
| CC | vectors, host cells and transgenic plants that are resistant to disease | |
| CC | are also provided. The fusion protein is target-specific, and resists | |
| CC | degradation in the insect gut | |
| XX | | |
| SQ | Sequence 554 AA; | |
| | Query Match 27.0%; Score 838.5; DB 6; Length 554; | |
| | Best Local Similarity 35.4%; Pred. No. 1.9e-52; | |
| | Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20; | |
| QY | 6 AILSLINACIGLWNASIKEDHNDYSKNPRIVCVGVGWSVYHK-VDPYTIEDIDPKCTHL 64 | |
| DB | 3 ATLATVALATAV-----QSDSRAPICVFVSNAVTRPGVGRYGIETIPVKCTHI 55 | |
| QY | 65 MYGFAKIDEYKTYTQVEDPYQDDNHNSWEKKGYERFNELRNKLNPETTMTISLGQVGESE 124 | |
| DB | 56 IYSFGVTGENSEVLIIDPELD-----VDNKGFRNTLSRSSHPSVKFWVAAGGAEGSS 110 | |
| QY | 125 KYSDMAANPTVRQDFTQSGLVDFTQEYKFDGLDLDWNYPGSR-LGNPKIDKQNYLAIVREL 183 | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:59:34 ; Search time 113.53 Seconds
(without alignments)
1265.926 Million cell updates/sec

Title: US-09-662-293-18

Perfect score: 3107
Sequence: 1 MKTIYAILSIMACIGLMNAS.....IMCPFGTICQEKLTIGE 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 3107 | 100.0 | 555 | 14 | US-10-218-743-15 |
| 2 | 3107 | 100.0 | 555 | 14 | US-10-218-743-18 |
| 3 | 3014 | 97.0 | 536 | 14 | US-10-218-743-21 |
| 4 | 2542 | 81.8 | 509 | 14 | US-10-218-743-35 |
| 5 | 2542 | 81.8 | 509 | 14 | US-10-218-743-38 |
| 6 | 2475 | 79.7 | 490 | 14 | US-10-218-743-41 |
| 7 | 738 | 23.8 | 445 | 14 | US-10-004-2198-10 |
| 8 | 738 | 23.8 | 466 | 14 | US-10-161-547-2 |
| 9 | 736.5 | 23.7 | 473 | 14 | US-10-004-2198-4 |
| 10 | 732 | 23.6 | 466 | 14 | US-10-161-547-4 |
| 11 | 729.5 | 23.5 | 452 | 14 | US-10-004-2198-9 |
| 12 | 721.5 | 23.2 | 373 | 14 | US-10-161-547-14 |
| 13 | 721.5 | 23.2 | 373 | 14 | US-10-161-547-15 |
| 14 | 708.5 | 22.8 | 455 | 14 | US-10-004-2198-14 |
| 15 | 708.5 | 22.8 | 476 | 14 | US-10-004-2198-1 |

| | | | | | | |
|----|-------|------|-----|----|----------------------|--------------------|
| 16 | 705.5 | 22.7 | 617 | 15 | US-10-369-493-6743 | Sequence 6743, Ap |
| 17 | 668 | 21.5 | 678 | 15 | US-10-295-027-6332 | Sequence 632, App |
| 18 | 662.5 | 21.3 | 457 | 15 | US-10-094-749-1900 | Sequence 1900, Ap |
| 19 | 662 | 21.3 | 383 | 14 | US-10-037-340-45 | Sequence 45, Appl |
| 20 | 662 | 21.3 | 383 | 15 | US-10-235-027-270 | Sequence 270, Appl |
| 21 | 650 | 20.9 | 383 | 9 | US-09-459-749D-17 | Sequence 17, Appl |
| 22 | 630 | 20.3 | 385 | 15 | US-10-295-027-278 | Sequence 278, App |
| 23 | 630 | 20.3 | 423 | 15 | US-10-373-802-2 | Sequence 2, Appli |
| 24 | 592 | 19.1 | 599 | 15 | US-10-369-493-17162 | Sequence 17162, A |
| 25 | 521.5 | 16.8 | 407 | 15 | US-10-369-493-20271 | Sequence 20271, A |
| 26 | 516 | 16.6 | 428 | 15 | US-10-369-493-16644 | Sequence 16644, A |
| 27 | 471 | 15.2 | 392 | 15 | US-10-369-493-19361 | Sequence 19361, A |
| 28 | 458 | 14.7 | 351 | 15 | US-10-369-493-12478 | Sequence 12478, A |
| 29 | 438 | 14.1 | 366 | 12 | US-10-424-599-209667 | Sequence 209667, A |
| 30 | 414.5 | 13.3 | 485 | 15 | US-10-369-493-3399 | Sequence 3399, Ap |
| 31 | 406 | 13.1 | 371 | 9 | US-09-923-844B-2 | Sequence 2, Appli |
| 32 | 405 | 13.0 | 546 | 15 | US-10-369-493-18584 | Sequence 18584, A |
| 33 | 401.5 | 12.9 | 399 | 14 | US-10-316-754-11 | Sequence 11, Appl |
| 34 | 401.5 | 12.9 | 399 | 14 | US-10-316-754-19 | Sequence 19, Appl |
| 35 | 372.5 | 12.0 | 449 | 15 | US-10-369-493-3308 | Sequence 3308, Ap |
| 36 | 367.5 | 11.8 | 511 | 15 | US-10-369-493-1684 | Sequence 1684, Ap |
| 37 | 352.5 | 11.3 | 572 | 15 | US-10-369-493-18585 | Sequence 18585, A |
| 38 | 350.5 | 11.3 | 358 | 12 | US-10-425-114-47084 | Sequence 47084, A |
| 39 | 350.5 | 11.3 | 365 | 12 | US-10-424-599-273219 | Sequence 273219, A |
| 40 | 347.5 | 11.2 | 567 | 15 | US-10-369-493-8158 | Sequence 8158, Ap |
| 41 | 337 | 10.8 | 789 | 15 | US-10-369-493-12804 | Sequence 12804, A |
| 42 | 329.5 | 10.6 | 445 | 15 | US-10-369-493-3089 | Sequence 3089, Ap |
| 43 | 317.5 | 10.2 | 462 | 15 | US-10-369-493-8244 | Sequence 8244, Ap |
| 44 | 314 | 10.1 | 432 | 15 | US-10-369-493-5521 | Sequence 5521, Ap |
| 45 | 313.5 | 10.1 | 607 | 14 | US-10-156-761-10415 | Sequence 10415, A |

ALIGNMENTS

RESULT 1

US-10-218-743-15
; Sequence 15, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-15

Query Match 100.0%; Score 3107; DB 14; Length 555;
Best Local Similarity 100.0%; Pred. No. 3e-220;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIYAILSIMACIGLMNASIKRDNDYSKQPMRIVCVGTWVYHKVDPTTIDIDPFK 60
Db 1 MKTIYAILSIMACIGLMNASIKRDNDYSKQPMRIVCVGTWVYHKVDPTTIDIDPFK 60

QY 61 CTHLMYGFPAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKLNKPELTTMISLGGWY 120
Db 61 CTHLMYGFPAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKLNKPELTTMISLGGWY 120
QY 121 EGSEKYSDMAANPTTRQOFIQSVLDFLOEYKFDGLDWEYFGSLGNPKIDKQNYLALV 180
Db 121 EGSEKYSDMAANPTTRQOFIQSVLDFLOEYKFDGLDWEYFGSLGNPKIDKQNYLALV 180
QY 181 RELKDAPEPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVMYDYHGGWENFYGHNA 240
Db 181 RELKDAPEPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVMYDYHGGWENFYGHNA 240
QY 241 PLYKRDEDELTHTYFNNVTMTHYLLNNGATRDKLVGMVFPFYGRAWSIEDRSKLGDPGA 300
Db 241 PLYKRDEDELTHTYFNNVTMTHYLLNNGATRDKLVGMVFPFYGRAWSIEDRSKLGDPGA 300
QY 301 KGMSPPGFIISGEGVLSYIELCOLFOKEEWHIQYDEYNNAPYGYNDKIWVGDDLASISC 360
Db 301 KGMSPPGFIISGEGVLSYIELCOLFOKEEWHIQYDEYNNAPYGYNDKIWVGDDLASISC 360
QY 361 KLAFLKELGSGVGMVWSLENDPFKHCGRPKNPLLNKVNHNMGDEKNSFECILGSPSTTTP 420
Db 361 KLAFLKELGSGVGMVWSLENDPFKHCGRPKNPLLNKVNHNMGDEKNSFECILGSPSTTTP 420
QY 421 TPSTTTTHTSETPKYTYVVDGHLIKCYKEGDIHPHTNIHKLVCFEFGVNGWVHIMPCP 540
Db 421 TPSTTTTHTSETPKYTYVVDGHLIKCYKEGDIHPHTNIHKLVCFEFGVNGWVHIMPCP 540
QY 541 PGTWCQEKLTICGE 555
Db 541 PGTWCQEKLTICGE 555

RESULT 2
US-10-218-743-18
; Sequence 18, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-18
Query Match 100.0%; Score 3107; DB 14; Length 555;
Best Local Similarity 100.0%; Pred. No. 3e-220;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTIYAILSIMACIGLMNASIKRDNDYSKNPMRIVCYVGTWVSVYHKVDPYTTIEDIDPPK 60
Db 1 MKTIYAILSIMACIGLMNASIKRDNDYSKNPMRIVCYVGTWVSVYHKVDPYTTIEDIDPPK 60
QY 61 CTHLMYGFPAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKLNKPELTTMISLGGWY 120
Db 61 CTHLMYGFPAKIDYKTYIQVDFPYQDDNHNSEKRGYERFNNLRKLNKPELTTMISLGGWY 120
QY 121 EGSEKYSDMAANPTTRQOFIQSVLDFLOEYKFDGLDWEYFGSLGNPKIDKQNYLALV 180
Db 121 EGSEKYSDMAANPTTRQOFIQSVLDFLOEYKFDGLDWEYFGSLGNPKIDKQNYLALV 180
QY 181 RELKDAPEPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVMYDYHGGWENFYGHNA 240
Db 181 RELKDAPEPHGYLLTAAVSPGKDKIDRAYDIKELNKLFDMMNVMYDYHGGWENFYGHNA 240
QY 241 PLYKRDEDELTHTYFNNVTMTHYLLNNGATRDKLVGMVFPFYGRAWSIEDRSKLGDPGA 300
Db 241 PLYKRDEDELTHTYFNNVTMTHYLLNNGATRDKLVGMVFPFYGRAWSIEDRSKLGDPGA 300
QY 301 KGMSPPGFIISGEGVLSYIELCOLFOKEEWHIQYDEYNNAPYGYNDKIWVGDDLASISC 360
Db 301 KGMSPPGFIISGEGVLSYIELCOLFOKEEWHIQYDEYNNAPYGYNDKIWVGDDLASISC 360
QY 361 KLAFLKELGSGVGMVWSLENDPFKHCGRPKNPLLNKVNHNMGDEKNSFECILGSPSTTTP 420
Db 361 KLAFLKELGSGVGMVWSLENDPFKHCGRPKNPLLNKVNHNMGDEKNSFECILGSPSTTTP 420
QY 421 TPSTTTTHTSETPKYTYVVDGHLIKCYKEGDIHPHTNIHKLVCFEFGVNGWVHIMPCP 540
Db 421 TPSTTTTHTSETPKYTYVVDGHLIKCYKEGDIHPHTNIHKLVCFEFGVNGWVHIMPCP 540
QY 541 PGTWCQEKLTICGE 555
Db 541 PGTWCQEKLTICGE 555

RESULT 3
US-10-218-743-21
; Sequence 21, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-21
Query Match 97.0%; Score 3014; DB 14; Length 536;

Best Local Similarity 100.0%; Pred. No. 2e-213; Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SIKRDNDYSKPMRIVCYVGTWSVYHKVDPTTIEDDPKCTHLMYGFADIKYKTIQ 79
Db 1 SIKRDNDYSKPMRIVCYVGTWSVYHKVDPTTIEDDPKCTHLMYGFADIKYKTIQ 60

QY 80 VFDPYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWYSEKYSMAANPTVROOF 139
Db 61 VFDPYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWYSEKYSMAANPTVROOF 120

QY 140 IQSVLDFLQBYKFDGLDWEYPSGRNLGNPKIDKQNYLALVRELKDAFPHGYLLTAAYS 199
Db 121 IQSVLDFLQBYKFDGLDWEYPSGRNLGNPKIDKQNYLALVRELKDAFPHGYLLTAAYS 180

QY 200 PGKDKIDRAYDIKELNKLFDMMNVMYDHYHGGWENFYGHNAPLYKRPDDELHTYFNVN 259
Db 181 PGKDKIDRAYDIKELNKLFDMMNVMYDHYHGGWENFYGHNAPLYKRPDDELHTYFNVN 240

QY 260 YTHVYLLNNGATRDKLVMGVFFYGRAWSIEDRSKLGDPAGKMSPPGFISGEEGVLSTY 319
Db 241 YTHVYLLNNGATRDKLVMGVFFYGRAWSIEDRSKLGDPAGKMSPPGFISGEEGVLSTY 300

QY 320 ELCQLFQKEBWHIYQDEYNNAPYGYNDKIWGYDDDLASISCKLAFKLGVSGVMWWSLE 379
Db 301 ELCQLFQKEBWHIYQDEYNNAPYGYNDKIWGYDDDLASISCKLAFKLGVSGVMWWSLE 360

QY 380 NDDFKHCGCPKPLNKNVNMINGDEKNSFECILGPSTTTPTTPTTPTTPTTPTTPT 439
Db 361 NDDFKHCGCPKPLNKNVNMINGDEKNSFECILGPSTTTPTTPTTPTTPTTPTTPT 420

QY 440 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
Db 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480

QY 500 VDGLHKYCKEKGDIHPHTNIHKYLVCFVNGWVHIMPCPGTIWCQKLCIG 555
Db 481 VDGLHKYCKEKGDIHPHTNIHKYLVCFVNGWVHIMPCPGTIWCQKLCIG 536

RESULT 4
US-10-218-743-35
; Sequence 35, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-35

Query Match 81.8%; Score 2542; DB 14; Length 509;
Best Local Similarity 81.7%; Pred. No. 1e-178;

Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

QY 1 MKTYAILISMACIGLMMNASIKRDNDYSKPMRIVCYVGTWSVYHKVDPTTIEDDPK 60
Db 1 MKTYPALFCIWACIGLMMNAATKRDNDYSKPMRIVCYVGTWSVYHKVDPTTIEDDPK 60

QY 61 CTHLMYGFADIKYKTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120
Db 61 CTHLMYGFADIKYKTIQVDFPDYQDDNHNSEKRGYERFNNLRKLNPELTMTISLGWY 120

QY 121 EGSEKYSMAANPTVROQFIQSVLDFLQBYKFDGLDWEYPSGRNLGNPKIDKQNYLALV 180
Db 121 EGSEKYSMAANPTVROQFIQSVLDFLQBYKFDGLDWEYPSGRNLGNPKIDKQNYLALV 180

QY 181 RELKDAFPHGYLLTAAYSPGKDKIDRAYDIKELNKLFDMMNVMYDHYHGGWENFYGHNA 240
Db 181 RELKDAFPHGYLLTAAYSPGKDKIDRAYDIKELNKLFDMMNVMYDHYHGGWENFYGHNA 240

QY 241 PLYKRPDDELHTYFNVNMYTHVYLLNNGATRDKLVMGVFFYGRAWSIEDRSKLGDP 300
Db 241 PLYKRPDDELHTYFNVNMYTHVYLLNNGATRDKLVMGVFFYGRAWSIEDRSKLGDP 300

QY 301 KGMSPPGFISGEEGVLSTYELCQLFQKEBWHIYQDEYNNAPYGYNDKIWGYDDDLASIS 360
Db 301 KGMSPPGFISGEEGVLSTYELCQLFQKEBWHIYQDEYNNAPYGYNDKIWGYDDDLASIS 360

QY 361 KLAFLKELGVSGVMWWSLENDDEKNGCPKPLNKNVNMINGDEKNSFECILGPSTTPT 420
Db 361 KLAFLKELGVSGVMWWSLENDDEKNGCPKPLNKNVNMINGDEKNSFECILGPSTTPT 420

QY 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
Db 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 434

QY 481 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 537
Db 435 ---PTTDDSTSETPKYTYIDGHLIKYCKEKGDIHPHTNIHKYLVCFV---NGWVHIM 491

QY 538 PCPPGTIWCQKLCIG 555
Db 492 DCPKTRHATFLKNCIOE 509

RESULT 5
US-10-218-743-38
; Sequence 38, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-38

Query Match 81.8%; Score 2542; DB 14; Length 509;
Best Local Similarity 81.7%; Pred. No. 1e-178; Indels 52; Gaps 2;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

1 MKTIVAILISACIGLMAASIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
1 MKTIFALFCIWACIGLMAAATKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
61 CTHLMYGFPAKIDYKTYIQVDPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWY 120
61 CTHLMYGFPAKIDYKTYIQVDPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWY 120
121 EGSEKYSMAANPTVROQFIOVSVDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALV 180
121 EGSEKYSMAANPTVROQFIOVSVDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALV 180
181 RELKDAEPHGVLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
181 RELKDAEPHGVLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
241 PLYKEPDETDLHTYFNNVNTYHYYLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPA 300
241 PLYKEPDETDLHTYFNNVNTYHYYLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPA 300
301 KMSPPGIFSGEGVLSVIELCOLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISC 360
301 KMSPPGIFSGEGVLSVIELCOLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISC 360
361 KLAFLKELGVSGVWVMSLENDFKHCGPKNPLKVNHNMGDEKNSFCILGSPSTTP 420
361 KLAFLKELGVSGVWVMSLENDFKHCGPKNPLKVNHNMGDEKNSFCILGSPSTTP 420
421 PTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
421 PTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
481 TPSPPTTHTSTPTKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIM 537
435 ---PTTDTSTSTPTKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIM 537
538 PCPPGTWICQELKTCIGE 555
492 DCPKGRWHTATLKNKICQE 509

RESULT 6
US-10-218-743-41
; Sequence 41, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490

TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-41

Query Match 79.7%; Score 2475; DB 14; Length 490;
Best Local Similarity 82.5%; Pred. No. 8.4e-174; Indels 52; Gaps 2;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;

22 KRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKTYIQV 81
3 KRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKTYIQV 82
82 DPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWYEGSEKYSMAANPTVROQFIO 141
63 DPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWYEGSEKYSMAANPTVROQFIO 122
142 SVLDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALVRELKDAEPHGVLLTAAYSPG 201
123 SVLDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALVRELKDAEPHGVLLTAAYSPG 182
202 KDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNNVNT 261
183 KDKIDVAYELKELNQLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNNVNT 242
262 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPAKMSPPGIFSGEGVLSVIEL 321
243 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPAKMSPPGIFSGEGVLSVIEL 302
322 COLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISCFLKELGVSGVWVMSLEND 381
303 COLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISCFLKELGVSGVWVMSLEND 362
382 DFHKGCHGPKNPLKVNHNMGDEKNSFCILGSPSTTPTTPTTPTTPTTPTTPTTPTT 441
363 DFHKGCHGPKNPLKVNHNMGDEKNSFCILGSPSTTPTTPTTPTTPTTPTTPTTPTT 415
442 TTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 501
416 -----PTTDTSTSTPTKYTYID 433
502 GHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIMPCPPGTWICQELKTCIGE 555
434 GHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIMPCPPGTWICQELKTCIGE 490

RESULT 7
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(445)
; OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match 23.8%; Score 738; DB 14; Length 445;
Best Local Similarity 34.2%; Pred. No. 5.3e-46;

Query Match 81.8%; Score 2542; DB 14; Length 509;
Best Local Similarity 81.7%; Pred. No. 1e-178; Indels 52; Gaps 2;
Matches 456; Conservative 24; Mismatches 26; Indels 52; Gaps 2;

1 MKTIVAILISACIGLMAASIKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
1 MKTIFALFCIWACIGLMAAATKRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFK 60
61 CTHLMYGFPAKIDYKTYIQVDPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWY 120
61 CTHLMYGFPAKIDYKTYIQVDPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWY 120
121 EGSEKYSMAANPTVROQFIOVSVDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALV 180
121 EGSEKYSMAANPTVROQFIOVSVDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALV 180
181 RELKDAEPHGVLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
181 RELKDAEPHGVLLTAAYSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
241 PLYKEPDETDLHTYFNNVNTYHYYLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPA 300
241 PLYKEPDETDLHTYFNNVNTYHYYLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPA 300
301 KMSPPGIFSGEGVLSVIELCOLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISC 360
301 KMSPPGIFSGEGVLSVIELCOLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISC 360
361 KLAFLKELGVSGVWVMSLENDFKHCGPKNPLKVNHNMGDEKNSFCILGSPSTTP 420
361 KLAFLKELGVSGVWVMSLENDFKHCGPKNPLKVNHNMGDEKNSFCILGSPSTTP 420
421 PTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
421 PTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
481 TPSPPTTHTSTPTKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIM 537
435 ---PTTDTSTSTPTKYTYVVDGHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIM 537
538 PCPPGTWICQELKTCIGE 555
492 DCPKGRWHTATLKNKICQE 509

RESULT 6
US-10-218-743-41
; Sequence 41, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490

TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-41

Query Match 79.7%; Score 2475; DB 14; Length 490;
Best Local Similarity 82.5%; Pred. No. 8.4e-174; Indels 52; Gaps 2;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;

22 KRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKTYIQV 81
3 KRDHNDYSKNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFPAKIDYKTYIQV 82
82 DPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWYEGSEKYSMAANPTVROQFIO 141
63 DPQDDNNHNSWKGVERFNNLRLKNPELTMTISLGWYEGSEKYSMAANPTVROQFIO 122
142 SVLDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALVRELKDAEPHGVLLTAAYSPG 201
123 SVLDLFOEYKFDGLDLDWEYPGSLGNPKDKQNYLALVRELKDAEPHGVLLTAAYSPG 182
202 KDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNNVNT 261
183 KDKIDVAYELKELNQLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNNVNT 242
262 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPAKMSPPGIFSGEGVLSVIEL 321
243 MHYLLNNGATRDKLVMGVPFYGRAWSIEDRSKVLGDPAKMSPPGIFSGEGVLSVIEL 302
322 COLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISCFLKELGVSGVWVMSLEND 381
303 COLFOKEEWHIOYDEYNNAPYGYNDKLVWGYDDIASISCFLKELGVSGVWVMSLEND 362
382 DFHKGCHGPKNPLKVNHNMGDEKNSFCILGSPSTTPTTPTTPTTPTTPTTPTTPTT 441
363 DFHKGCHGPKNPLKVNHNMGDEKNSFCILGSPSTTPTTPTTPTTPTTPTTPTTPTT 415
442 TTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 501
416 -----PTTDTSTSTPTKYTYID 433
502 GHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIMPCPPGTWICQELKTCIGE 555
434 GHLIKCYKEGDIHPHTNIHKYLVCEV---NGGWWVHIMPCPPGTWICQELKTCIGE 490

RESULT 7
US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(445)
; OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match 23.8%; Score 738; DB 14; Length 445;
Best Local Similarity 34.2%; Pred. No. 5.3e-46;

| | | | | | | | | | |
|---------|------|--------------|-----|------------|------|--------|-----|------|-----|
| Matches | 152; | Conservative | 85; | Mismatches | 150; | Indels | 58; | Gaps | 11; |
|---------|------|--------------|-----|------------|------|--------|-----|------|-----|

| | | | |
|----|-----|--|-----|
| QY | 34 | RIVCYVGTWSVYHKVDP--YTIEDIDDPFKCTHLMYGFAKIDEBYKYTIQVDFPDQDDNNHNSW | 92 |
| Db | 2 | KLVCYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGMTNHQLST-----TEW | 50 |
| QY | 93 | -EKXGYERFNNLRKLNPELTMISLGGWYEGSEKYSDMAANPTVYRQFIQSVLDFLOEYK | 151 |
| Db | 51 | NDETLYQEFNGLKQGNPKLTKLLAIGGWNFGTQFTDMVATANNRQTFVNSAIRFLRKY | 110 |
| QY | 152 | FDGLDLWEYVPGSRLGNPKIDKQNYLAIVRELKDAFEPHG-----YLLTAAVSPGDK | 204 |
| Db | 111 | FDGLDLWEYVPGSQ--GSPAVDKERTFTLVQDLANAFQOEATQSGKERILLISAAVPAQOTY | 169 |
| QY | 205 | IDRAYDIKELNKLDFWNNVMTYDYHGGWENFYGHNAPLYKEPDETDELHTVFNNTYTMHY | 264 |
| Db | 170 | VDAGYEVDKLAQNLDVFNLMAYDFHGSWEKVTGHNPSLYKQOESGGAAS--LNYDAAVQQ | 228 |
| QY | 265 | YLNNGATRDKLVMGVPPVYGRAWSIEDRSKLGKLPAPKMSPPGTSIGEEGVLSTIELCOL | 322 |
| Db | 229 | WLQKGTSPASKULLGNPNFYGRSFTLIASSSDTRVGAPATGSGTPGFTKEGMLAYEVCWS | 288 |
| QY | 325 | FQKEEWHIQDEYNNAPYGYNDKIWGVYDDLASISCKLAFELKEILGVSGVMVWSLENDDFK | 384 |
| Db | 289 | KGATKQRIQ-DQ--KVPIYPRDNQWGFDDVSEFKTKVSYLKQKGLGAMVWALDLDFA | 345 |
| QY | 385 | GHGCPKNPLLNKVNMMINGDEKNSEFICLSPSTTTPTTTPTTTPTTTPTTTPTTPTTPTT | 444 |
| Db | 346 | G-----FSCNQG-----RVPLIQLTRQELSLPVLPSGT | 373 |
| QY | 445 | PSPTPTTPTTSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT | 469 |
| Db | 374 | PELEVP-KPGQPSPEFHGPGSPGQDT | 397 |


```

RESULT 8
US-10-161-547-2
; Sequence 2, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Thjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITTIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-2

```


| | | | | | | | |
|-----------------------|--------|--------------|-------------|------------|------|--------|------|
| Query Match | 23.8%; | Score | 738; | DB | 14; | Length | 466; |
| Best Local Similarity | 34.2%; | Pred. | No.5.6e-46; | | | | |
| Matches | 152; | Conservative | 85; | Mismatches | 150; | Indels | 58; |
| Gaps | 11; | | | | | | |

| | | | |
|----|-----|--|-----|
| QY | 34 | RIVCYVGTWSVYHKVDP--YTIEDIDDPFKCTHLMYGFAKIDEBYKYTIQVDFPDQDDNNHNSW | 92 |
| Db | 23 | KLVCYFTNWAQYRQGEARFLPKDLDPSLCTHLIYAFAGMTNHQLST-----TEW | 71 |
| QY | 93 | -EKXGYERFNNLRKLNPELTMISLGGWYEGSEKYSDMAANPTVYRQFIQSVLDFLOEYK | 151 |
| Db | 72 | NDETLYQEFNGLKQGNPKLTKLLAIGGWNFGTQFTDMVATANNRQTFVNSAIRFLRKY | 131 |
| QY | 152 | FDGLDLWEYVPGSRLGNPKIDKQNYLAIVRELKDAFEPHG-----YLLTAAVSPGDK | 204 |
| Db | 132 | FDGLDLWEYVPGSQ--GSPAVDKERTFTLVQDLANAFQOEATQSGKERILLISAAVPAQOTY | 190 |


```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-14

Query Match      23.2%; Score 721.5; DB 14; Length 373;
Best Local Similarity 37.7%; Pred. No. 6.9e-45;
Matches 136; Conservative 83; Mismatches 117; Indels 25; Gaps 8;

QY 34 RIVCYVGTWSVYHKVDP-YTIEDIDPFKCTHLMYGFADKIDVYKTIQVDPYQDDNHSW 92
Db 2 KLVCYFTNWAQYRQGEARFPLKDLPSLCTHLIYAFAGMTNHQLSLST-----TEW 50
QY 93 -EKRGYERFNNLRKLNKPELTMTISLGGWYEGSEKYSMDAANPYRQOFTQSVLDLQYK 151
Db 51 NDETLVQEFNGLKQNPCKLTKLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYS 110
QY 152 FDGLDLWEYPGSRIGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASVPGKDK 204
Db 111 FDGLDLWEYPGSQ-GSPAVDKERFTTLVQDLANAFQOEAGTSGKERILLLSAAVPAGQTY 169
QY 205 IDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAFLYKRPDDELTHTYFNVTMWHY 264
Db 170 VDAGYEVDKIAQNLDVFNLMAYDFHGSWEKVTGHSNPLYKQESGAAS-LNVDAAVQO 228
QY 265 YLNGATRDKLVMGVFFYGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYELCOL 324
Db 229 WLQGTGTPASKILGMPTVGRSFTLASSSDTRVGAPATGSGTGPFTKEGGLMAYEVCWS 288
QY 325 FQKEEWHIQVDEYNAPYGYNDKIWVGYYDILASISCKLAFKLKELGVSVMWSLENDDFK 384
Db 289 KGATKQRIQ-DQ--KVPYIFRDNQWVGDFDVESEKTKVSYLKQKGLGGAMWALDLDFA 345
QY 385 G 385
Db 346 G 346

RESULT 14
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14

Query Match      22.8%; Score 708.5; DB 14; Length 455;
Best Local Similarity 35.9%; Pred. No. 8.1e-44;
Matches 149; Conservative 75; Mismatches 154; Indels 37; Gaps 10;

QY 34 RIVCYVGTWSVYHK-VDPYTIETIDPFKCTHLMYGFADKIDVYKTIQVDPYQDDNHSW 92
Db 2 QLTCTYTNWAQYRPGLGFRFMDNIDPCLTCLHLIYAFAGRONNEIT-----TEW 50
QY 93 -EKRGYERFNNLRKLNKPELTMTISLGGWYEGSEKYSMDAANPYRQOFTQSVLDLQYK 151
Db 51 NDETLVQEFNGLKQNPCKLTKLLAIGGNWFGTKFTDMVATANNRQTFVNSAIRFLRKYS 110
QY 152 FDGLDLWEYPGSRIGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASVPGKDK 204
Db 111 FDGLDLWEYPGSR-GSPPDQKHLFTVLVQEMREAEQKQINPKPLMTAAVAAGISN 169
QY 205 IDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAFLYKRPDDELTHTYFNVTMWHY 264
Db 170 IQSGYEIPQLSQYLDYIHMVTYDLHGSWEVGTGENSPLYKYPTDTGS-NAYLNDVYVMNY 228
QY 265 YLNGATRDKLVMGVFFYGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYELCOL 324
Db 229 WKDNGAPAEKLVGPFPTGHNFILNSPNTGTGAGTSGAGPAGYAKSGGIWAYEIC-T 287
```

```
QY 325 FQKEWHIQDEYYNAPYGVNDKIWGVYDDIASISCKLAFKLKELGVSGVMVWSLENDDEFK 384
Db 288 FLKNGATCGWDAPQGVYAYQGVNFWGVYDNKSFDDIQAQWLKHNKFCGANVWALDDDDFT 347
QY 385 G-HCGP-KNPLLNKVHNMINGDEKNSPECILGPSTTTPTTPTTPTTPTTPTTPTPS 437
Db 348 GTFCNQGKFPLI-----STLKKALGLQASCTAPAQPIETITAAPSGS 390

RESULT 15
US-10-004-219B-1
; Sequence 1, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2193-5136US
; CURRENT APPLICATION NUMBER: US/10/004.219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence
US-10-004-219B-1

Query Match 22.8%; Score 708.5; DB 14; Length 476;
Best Local Similarity 35.9%; Pred. No. 8.6e-44;
Matches 149; Conservative 75; Mismatches 154; Indels 37; Gaps 10;

QY 34 RIVCYGVGWSVVHK-VDPYTIEDIIDPFKCTHLMYGFAKIDBYKYTIQVDFPYQDDNHSW 92
Db 23 QLTCTFTWQAQYRPGUGRFMPDNIDPCLCTHLIYAFAGRQNEITT-----IEW 71
QY 93 -EKRGYERFNNLRNLKPELTMTMISLGGWYEGSEKYSMDAANPTYRQQFIQSVDLFLQYK 151
Db 72 NDTVLYQAFNGLKNKNSQLKTLAIGGNWFGTAPTAMVSTPENRQTEITSVIXFLRQYE 131
QY 152 FDGLDLDWEYPCSRIGNPKIDKQNYALVRELKDAPEPHG-----YLLTAAYSPGKOK 204
Db 132 FDGLDFDWEYPGSR-GSPPDQKHLFTVLVQEMREAFQEAQINKPRLMWTAAVAAGISN 190
QY 205 IDRAYDIKELNKLFDWMVMTYDHYHGWENFYGHNAPLYKRPDETDELHTYFNVNYTMHY 264
Db 191 IQSGYEIPQLSOYLDYIHVMTYDLHGSWEGYTGENSEPLYKYPTDTGS-NAYLNVDYVWNY 249
QY 265 YLNNGATRDKLVMGVPPFYGRAMSIEDRSKLGDPKAGMSPPGFTSGEGVLSYIELCOL 324
Db 250 WKDNGAPAEKLVGPTTYGHNFILNPSNTGIGAPTSAGPAGPYAKESGIWAYYEIC-T 308
QY 325 FQKEWHIQDEYYNAPYGVNDKIWGVYDDIASISCKLAFKLKELGVSGVMVWSLENDDEFK 384
Db 309 FLKNGATCGWDAPQGVYAYQGVNFWGVYDNKSFDDIQAQWLKHNKFCGANVWALDDDDFT 368
QY 385 G-HCGP-KNPLLNKVHNMINGDEKNSPECILGPSTTTPTTPTTPTTPTTPTTPTPS 437
Db 369 GTFCNQGKFPLI-----STLKKALGLQASCTAPAQPIETITAAPSGS 411
```

Search completed: March 22, 2004, 07:45:52
Job time : 114.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: March 22, 2004, 06:59:34 ; Search time 109.643 Seconds
(without alignments)
1265.926 Million cell updates/sec

Title: US-09-662-293-21

Perfect score: 3014

Sequence: 1 SIKRDNYSKPMRIVCVV.....IMPCPGTTCWEKLTICGE 536

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 25895339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 3014 | 100.0 | 536 | 14 | US-10-218-743-21 |
| 2 | 3014 | 100.0 | 555 | 14 | US-10-218-743-15 |
| 3 | 3014 | 100.0 | 555 | 14 | US-10-218-743-18 |
| 4 | 2475 | 82.1 | 490 | 14 | US-10-218-743-41 |
| 5 | 2475 | 82.1 | 509 | 14 | US-10-218-743-35 |
| 6 | 2475 | 82.1 | 509 | 14 | US-10-218-743-38 |
| 7 | 738 | 24.5 | 445 | 14 | US-10-004-219B-10 |
| 8 | 738 | 24.5 | 466 | 14 | US-10-161-547-2 |
| 9 | 732 | 24.3 | 465 | 14 | US-10-161-547-4 |
| 10 | 729.5 | 24.2 | 452 | 14 | US-10-004-219B-9 |
| 11 | 729.5 | 24.2 | 473 | 14 | US-10-004-219B-4 |
| 12 | 721.5 | 23.9 | 373 | 14 | US-10-161-547-14 |
| 13 | 721.5 | 23.9 | 373 | 14 | US-10-161-547-15 |
| 14 | 708.5 | 23.5 | 455 | 14 | US-10-004-219B-14 |
| 15 | 708.5 | 23.5 | 476 | 14 | US-10-004-219B-1 |

| | | | | | | |
|----|-------|------|-----|----|----------------------|-------------------|
| 16 | 705.5 | 23.4 | 617 | 15 | US-10-369-493-6743 | Sequence 6743, Ap |
| 17 | 667.5 | 22.1 | 678 | 15 | US-10-295-027-632 | Sequence 632, Ap |
| 18 | 662.5 | 22.0 | 457 | 15 | US-10-094-749-1900 | Sequence 1900, Ap |
| 19 | 660 | 21.9 | 383 | 14 | US-10-097-340-45 | Sequence 45, Appl |
| 20 | 660 | 21.9 | 383 | 15 | US-10-295-027-270 | Sequence 270, Ap |
| 21 | 650 | 21.6 | 383 | 9 | US-09-459-749D-17 | Sequence 17, Appl |
| 22 | 630 | 20.9 | 385 | 15 | US-10-295-027-278 | Sequence 278, Ap |
| 23 | 630 | 20.9 | 423 | 15 | US-10-373-802-2 | Sequence 2, Appli |
| 24 | 588 | 19.5 | 599 | 15 | US-10-369-493-17162 | Sequence 17162, A |
| 25 | 521.5 | 17.3 | 407 | 15 | US-10-369-493-20271 | Sequence 20271, A |
| 26 | 516 | 17.1 | 428 | 15 | US-10-369-493-16644 | Sequence 16644, A |
| 27 | 471 | 15.6 | 392 | 15 | US-10-369-493-19361 | Sequence 19361, A |
| 28 | 458 | 15.2 | 351 | 15 | US-10-369-493-12478 | Sequence 12478, A |
| 29 | 434.5 | 14.4 | 366 | 12 | US-10-424-599-209667 | Sequence 209667 |
| 30 | 414.5 | 13.8 | 485 | 15 | US-10-369-493-3399 | Sequence 3399, Ap |
| 31 | 406 | 13.5 | 371 | 9 | US-09-923-844B-2 | Sequence 2, Appli |
| 32 | 405 | 13.4 | 546 | 15 | US-10-369-493-18584 | Sequence 18584, A |
| 33 | 401.5 | 13.3 | 399 | 14 | US-10-316-754-11 | Sequence 11, Appl |
| 34 | 401.5 | 13.3 | 399 | 14 | US-10-316-754-19 | Sequence 19, Appl |
| 35 | 372.5 | 12.4 | 449 | 15 | US-10-369-493-3308 | Sequence 3308, Ap |
| 36 | 367.5 | 12.2 | 511 | 15 | US-10-369-493-1684 | Sequence 1684, Ap |
| 37 | 352.5 | 11.7 | 572 | 15 | US-10-369-493-18585 | Sequence 18585, A |
| 38 | 350.5 | 11.6 | 358 | 12 | US-10-425-114-47084 | Sequence 47084, A |
| 39 | 350.5 | 11.6 | 365 | 12 | US-10-424-599-273219 | Sequence 273219 |
| 40 | 347.5 | 11.5 | 567 | 15 | US-10-369-493-8158 | Sequence 8158, Ap |
| 41 | 337 | 11.2 | 789 | 15 | US-10-369-493-12804 | Sequence 12804, A |
| 42 | 329.5 | 10.9 | 445 | 15 | US-10-369-493-3089 | Sequence 3089, Ap |
| 43 | 317.5 | 10.5 | 462 | 15 | US-10-369-493-8244 | Sequence 8244, Ap |
| 44 | 313.5 | 10.4 | 607 | 14 | US-10-156-761-10415 | Sequence 10415, A |
| 45 | 311.5 | 10.3 | 427 | 14 | US-10-156-761-14446 | Sequence 14446, A |

ALIGNMENTS

RESULT 1

US-10-218-743-21
; Sequence 21, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-21

Query Match 100.0%; Score 3014; DB 14; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.2e-212;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SIKRDNYSKPMRIVCVVGTWSYHKVDPTTIEDIDPFKCTHLMYGFAKIDEXYTIQ 60
DB 1 SIKRDNYSKPMRIVCVVGTWSYHKVDPTTIEDIDPFKCTHLMYGFAKIDEXYTIQ 60

QY 61 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGWYEGSEKYSMDMAANPTYRQOF 120
 DB 61 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGWYEGSEKYSMDMAANPTYRQOF 120
 QY 121 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTAAS 180
 DB 121 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTAAS 180
 QY 181 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNNV 240
 DB 181 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNNV 240
 QY 241 YTMHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPKAGMSPPGFIISGEEGLVSYI 300
 DB 241 YTMHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPKAGMSPPGFIISGEEGLVSYI 300
 QY 301 ELCQLFQKEEWHIQDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVMWMSLE 360
 DB 301 ELCQLFQKEEWHIQDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVMWMSLE 360
 QY 361 NDDFKHCGCPNPLNKNVNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTTPTTPTTPT 420
 DB 361 NDDFKHCGCPNPLNKNVNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTTPTTPTTPT 420
 QY 421 TPT 480
 DB 421 TPT 480
 QY 481 VDGLIKCYKEGDIHPHTNIHKLVCFVNGWVWVHIMPCPGTIWCQEKLTICIG 536
 DB 481 VDGLIKCYKEGDIHPHTNIHKLVCFVNGWVWVHIMPCPGTIWCQEKLTICIG 536

RESULT 2
 US-10-218-743-15
 ; Sequence 15, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; PRIOR FILING DATE: 2002-08-13
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 555
 ; TYPE: PRT
 ; ORGANISM: Dermatophagoides farinae
 US-10-218-743-15

Query Match 100.0%; Score 3014; DB 14; Length 555;
 Best Local Similarity 100.0%; Pred. No. 1.3e-212;
 Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SIKRDNDYKSNPMRIVCVYVGTWSVHKVDPTTIEDIDPFCKTHLMYGFADIKYKTIQ 60
 DB 20 SIKRDNDYKSNPMRIVCVYVGTWSVHKVDPTTIEDIDPFCKTHLMYGFADIKYKTIQ 79

QY 61 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGWYEGSEKYSMDMAANPTYRQOF 120
 DB 80 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGWYEGSEKYSMDMAANPTYRQOF 139
 QY 121 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTAAS 180
 DB 140 IQSVLDFLOEYKFDGLDLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGVLTAAS 199
 QY 181 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNNV 240
 DB 200 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNNV 259
 QY 241 YTMHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPKAGMSPPGFIISGEEGLVSYI 300
 DB 260 YTMHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPKAGMSPPGFIISGEEGLVSYI 319
 QY 301 ELCQLFQKEEWHIQDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVMWMSLE 360
 DB 320 ELCQLFQKEEWHIQDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVMWMSLE 379
 QY 361 NDDFKHCGCPNPLNKNVNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTTPTTPTTPT 420
 DB 380 NDDFKHCGCPNPLNKNVNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTTPTTPTTPT 439
 QY 421 TPT 480
 DB 440 TPT 499
 QY 481 VDGLIKCYKEGDIHPHTNIHKLVCFVNGWVWVHIMPCPGTIWCQEKLTICIG 536
 DB 500 VDGLIKCYKEGDIHPHTNIHKLVCFVNGWVWVHIMPCPGTIWCQEKLTICIG 555

RESULT 3
 US-10-218-743-18
 ; Sequence 18, Application US/10218743
 ; Publication No. US20030096779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/10/218,743
 ; PRIOR FILING DATE: 2002-08-13
 ; PRIOR FILING DATE: 1999-04-15
 ; PRIOR FILING DATE: 1998-09-02
 ; PRIOR FILING DATE: 1998-05-13
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR FILING DATE: 1998-04-17
 ; PRIOR FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 18
 ; LENGTH: 555
 ; TYPE: PRT
 ; ORGANISM: Dermatophagoides farinae
 US-10-218-743-18

Query Match 100.0%; Score 3014; DB 14; Length 555;
 Best Local Similarity 100.0%; Pred. No. 1.3e-212;
 Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SIKRDNDYKSNPMRIVCVYVGTWSVHKVDPTTIEDIDPFCKTHLMYGFADIKYKTIQ 60
 DB 20 SIKRDNDYKSNPMRIVCVYVGTWSVHKVDPTTIEDIDPFCKTHLMYGFADIKYKTIQ 79
 QY 61 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGWYEGSEKYSMDMAANPTYRQOF 120

Db 80 VFDPYQDDNHNWSEKGYERFNNLRKLNPELTTMISLGWYEGSEKYSDMAANPTYRQOF 139
Qy 121 IQSVLDFOLEYKFDGLDWEYFGSLGNPKIDKQNYLALVRELKDAEPHGYLLTAAYS 180
Db 140 IQSVLDFOLEYKFDGLDWEYFGSLGNPKIDKQNYLALVRELKDAEPHGYLLTAAYS 199
Qy 181 PGKDKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNAPLYKRPDTEDELHITYFNN 240
Db 200 PGKDKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNAPLYKRPDTEDELHITYFNN 259
Qy 241 YTHYYLNGATRDKLWGVPPFYGRAWSIEDRSKLGDPAGKMPGPGFISGEEGVL 300
Db 260 YTHYYLNGATRDKLWGVPPFYGRAWSIEDRSKLGDPAGKMPGPGFISGEEGVL 319
Qy 301 ELQOLPQKEBWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGVSGVMWSLE 360
Db 320 ELQOLPQKEBWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGVSGVMWSLE 379
Qy 361 NDDPKHCHGCPKPLNKNVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPT 420
Db 380 NDDPKHCHGCPKPLNKNVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPT 439
Qy 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
Db 440 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
Qy 481 VDGLIKCYKEGDIHPHTNIHKYLVEFVNGWVHIMCPPGTIWCQEKLCIG 536
Db 500 VDGLIKCYKEGDIHPHTNIHKYLVEFVNGWVHIMCPPGTIWCQEKLCIG 555

RESULT 4
US-10-218-743-41
; Sequence 41, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-41

Query Match 82.1%; Score 2475; DB 14; Length 490;
Best Local Similarity 82.5%; Pred. No. 3.6e-173;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
Qy 3 KRDNHNSKPMRIVCYVGTWSVYHKVDPTIEDIDPFKCTHLMYGFAKIDYKTIQVF 62
Db 3 KRDNHNSKPMRIVCYVGTWSVYHKVDPTIEDIDPFKCTHLMYGFAKIDYKTIQVF 62
Qy 63 DPYQDDNHNWSEKGYERFNNLRKLNPELTTMISLGWYEGSEKYSDMAANPTYRQOF 122

Db 63 DPFODDNHNWSEKGYERFNNLRKLNPELTTMISLGWYEGSEKYSDMAANPTYRQOF 122
Qy 123 SVLDFOLEYKFDGLDWEYFGSLGNPKIDKQNYLALVRELKDAEPHGYLLTAAYS 182
Db 123 SVLDFOLEYKFDGLDWEYFGSLGNPKIDKQNYLALVRELKDAEPHGYLLTAAYS 182
Qy 183 KDKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNAPLYKRPDTEDELHITYFNN 242
Db 183 KDKIDRAYDIKELNKLFDWMNMTYDYGWENFYGHNAPLYKRPDTEDELHITYFNN 242
Qy 243 MYHYLNGATRDKLWGVPPFYGRAWSIEDRSKLGDPAGKMPGPGFISGEEGVL 302
Db 243 MYHYLNGATRDKLWGVPPFYGRAWSIEDRSKLGDPAGKMPGPGFISGEEGVL 302
Qy 303 COLPQKEBWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGVSGVMWSLE 362
Db 303 COLPQKEBWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGVSGVMWSLE 362
Qy 363 DFKGHCPCPKPLNKNVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPT 422
Db 363 DFKGHCPCPKPLNKNVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPT 415
Qy 423 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 482
Db 416 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 433
Qy 483 GHLIKCYKEGDIHPHTNIHKYLVEFVNGWVHIMCPPGTIWCQEKLCIG 536
Db 434 GHLIKCYKEGDIHPHTNIHKYLVEFVNGWVHIMCPPGTIWCQEKLCIG 490

RESULT 5
US-10-218-743-35
; Sequence 35, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-35

Query Match 82.1%; Score 2475; DB 14; Length 509;
Best Local Similarity 82.5%; Pred. No. 3.8e-173;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
Qy 3 KRDNHNSKPMRIVCYVGTWSVYHKVDPTIEDIDPFKCTHLMYGFAKIDYKTIQVF 62
Db 22 KRDNHNSKPMRIVCYVGTWSVYHKVDPTIEDIDPFKCTHLMYGFAKIDYKTIQVF 81
Qy 63 DPYQDDNHNWSEKGYERFNNLRKLNPELTTMISLGWYEGSEKYSDMAANPTYRQOF 122
Db 82 DPFODDNHNWSEKGYERFNNLRKLNPELTTMISLGWYEGSEKYSDMAANPTYRQOF 141

QY 123 SVLDFLOEYKFDGLDLDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPG 182
DB 142 SVLDFLOEYKFDGLDLDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPG 201
QY 183 KDKIDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNVT 242
DB 202 KDKIDVAYELKELNQLFDWMNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNVT 261
QY 243 MHYILNNGATRDKLVMGVFFYGRAWSIEDRSKVLGDPKAGMSPPGFISGEGVLSIEL 302
DB 262 MHYILNNGATRDKLVMGVFFYGRAWSIEDRSKVLGDPKAGMSPPGFISGEGVLSIEL 321
QY 303 COLFOKEEWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGELGVGVWVMSLEND 362
DB 322 COLFOKEEWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGELGVGVWVMSLEND 381
QY 363 DFKHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTTPTTTP 422
DB 382 DFKHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTTPTT 434
QY 423 TTTTPTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 482
DB 435 -----PTTDTSTSETPKYTYID 452
QY 483 GHLIKYKEDIPHPNTNHHKYLVCFFV---NGGWWHIMPCCPGTIWCEKLTICGE 536
DB 453 GHLIKYKQGYLPHPTDVHKYLVCYIATPNGGWWHIMDCPKGTRWHATLKNKCIQE 509

RESULT 6

US-10-218-743-38
; Sequence 38, Application US/10218743
; Publication No. US20030096779A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Huster, Shirley W.
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/10/218,743
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/09/292,225
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 60/098,909
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/085,295
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,565
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: 09/062,013
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 38
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-10-218-743-38

Query Match 82.1%; Score 2475; DB 14; Length 509;
Best Local Similarity 82.5%; Pred. No. 3.8e-173;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
QY 3 KRDNHNSKPNRIVCYVGTWSVYHKVDPTIEDIPFKCTHLMYGFADIKYTIQVF 62
DB 22 KRDNHNSKPNRIVCYVGTWSVYHKVDPTIEDIPFKCTHLMYGFADIKYTIQVF 81
QY 63 DPODDNHNSEKGYERFNNLRKNPELTMTISLGGWYEGSEKYSMAANPTVROQFIO 122
DB 82 DPODDNHNSEKGYERFNNLRKNPELTMTISLGGWYEGSEKYSMAANPTVROQFIO 141

QY 123 SVLDFLOEYKFDGLDLDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPG 182
DB 142 SVLDFLOEYKFDGLDLDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPG 201
QY 183 KDKIDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNVT 242
DB 202 KDKIDVAYELKELNQLFDWMNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNVT 261
QY 243 MHYILNNGATRDKLVMGVFFYGRAWSIEDRSKVLGDPKAGMSPPGFISGEGVLSIEL 302
DB 262 MHYILNNGATRDKLVMGVFFYGRAWSIEDRSKVLGDPKAGMSPPGFISGEGVLSIEL 321
QY 303 COLFOKEEWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGELGVGVWVMSLEND 362
DB 322 COLFOKEEWHIOYDEYNAPYGYNDKIWGVYDDLASISCKLAFKELGELGVGVWVMSLEND 381
QY 363 DFKHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTTPTTTP 422
DB 382 DFKHCGKPNLLKNVHNMINGDEKNSPECILGSPSTTTPTTTPTTTPTTTPTT 434
QY 423 TTTTPTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 482
DB 435 -----PTTDTSTSETPKYTYID 452
QY 483 GHLIKYKEDIPHPNTNHHKYLVCFFV---NGGWWHIMPCCPGTIWCEKLTICGE 536
DB 453 GHLIKYKQGYLPHPTDVHKYLVCYIATPNGGWWHIMDCPKGTRWHATLKNKCIQE 509

RESULT 7

US-10-004-219B-10
; Sequence 10, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1) (445)
; OTHER INFORMATION: /note="Human chitotriosidase"
US-10-004-219B-10

Query Match 24.5%; Score 738; DB 14; Length 445;
Best Local Similarity 34.2%; Pred. No. 7.9e-46;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;
QY 15 RIVCYVGTWSVYHKVDPTIEDIPFKCTHLMYGFADIKYTIQVDPDQDDNHNWS 73
DB 2 KLVCFYTNNAQYRQGEARELPKDLDPSCVTHLYAFAGMTNHLST-----TGW 50
QY 74 -EKGYERFNNLRKNPELTMTISLGGWYEGSEKYSMAANPTVROQFISVLDFOEYK 132
DB 51 NDELTQYENGLKKNPKLTLLAIGGWNFGQKFTDMVATANNQTFVNSAIRLRKYS 110
QY 133 FUGLDDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPGKDK 185
DB 111 FUGLDDWEYFSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAASPGKDK 169
QY 186 IDRAYDIKELNKLFDWMNVMTYDYGWENFYGHNAPLYKRPDETDELHTYFNVT 245

Db 170 VDAGYEVDKIAQNLDVNLVAYDFHGSWEKVTGHNSPLYKREQESGAAS-LNVDAAVQ 228
Qy 246 YLNGATRDKLVMPVPGYGRWSIEDRSKLGDPKAGMSPPGFISGEEGVLSYIELCQL 305
Db 229 WLQKGTTPASKLILGMPTYSFSLASSSDTRVGAPATGSGTGPFTKEGGLMAYEVCWS 288
Qy 306 FQKEEWHIOYDEYVYNAFYNDKIWGYDDLASISCKLAFKELGVSVMVMSLENDDEFK 365
Db 289 KGATKQRIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSKLQKGLGAMVWALDDFA 345
Qy 366 GHCGPKNPLLNKVNHMGDEKNSFECILGPGSTTTPTTPTTPTTPTTPTTPTTPTTPTT 425
Db 346 G-----FSCNQ-----RYPLIQTLRQELSPLPLPSGT 373
Qy 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450
Db 374 PELEVP-KPGQSPSEPHGSPGQDT 397

RESULT 8
US-10-161-547-2
; Sequence 2, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/039,198
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-2

Query Match 24.5%; Score 738; DB 14; Length 466;
Best Local Similarity 34.2%; Pred. No. 8.3e-46;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;

Qy 15 RIVCYGTWSVYHKVDP-YTIEDIDPFKCTHLMYGFADKIDYKTIQVDFPYQDDNHSW 73
Db 23 KLVCFYTNWQYRQGEARFLPKOLDPSLCTHLIYAFAGMTNHQLST-----TEW 71
Qy 74 -EKGYERFNNLRKNPELTMTISLGGWYEGSEKYSDMAANPTYRQFIQSVDLFLQYK 132
Db 72 NDETLYOEFNGLKKNPKLTLAIGWNFTQKFTDMVATANNRQTFVNSAIRLKYK 131
Qy 133 FDGLDLWEYPGSRIGNPKDKONYALVRELKDAFEPHG-----YLLTAASVPGDK 185
Db 132 FDGLDLWEYPGSQ--GSPAVDKERTTLVQDLANAFQOEAOQSGKERLLLSAAVPAGQY 190
Qy 186 IDRAYDIKELNKLFDWNVMTYDYGWENFYGHNAPLYKRPDDELTHTYFNVTMTHY 245
Db 191 VDAGYEVDKIAQNLDVNLVAYDFHGSWEKVTGHNSPLYKREQESGAAS-LNVDAAVQ 249
Qy 246 YLNGATRDKLVMPVPGYGRWSIEDRSKLGDPKAGMSPPGFISGEEGVLSYIELCQL 305
Db 250 WLQKGTTPASKLILGMPTYSFSLASSSDTRVGAPATGSGTGPFTKEGGLMAYEVCWS 309
Qy 306 FQKEEWHIOYDEYVYNAFYNDKIWGYDDLASISCKLAFKELGVSVMVMSLENDDEFK 365
Db 310 KGATKQRIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSKLQKGLGAMVWALDDFA 366
Qy 366 GHCGPKNPLLNKVNHMGDEKNSFECILGPGSTTTPTTPTTPTTPTTPTTPTTPTTPTT 425
Db 367 G-----FSCNQ-----RYPLIQTLRQELSPLPLPSGT 394

Qy 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450
Db 395 PELEVP-KPGQSPSEPHGSPGQDT 418

RESULT 9
US-10-161-547-4
; Sequence 4, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-4

Query Match 24.3%; Score 732; DB 14; Length 466;
Best Local Similarity 33.9%; Pred. No. 2.3e-45;
Matches 151; Conservative 85; Mismatches 151; Indels 58; Gaps 11;

Qy 15 RIVCYGTWSVYHKVDP-YTIEDIDPFKCTHLMYGFADKIDYKTIQVDFPYQDDNHSW 73
Db 23 KLVCFYTNWQYRQGEARFLPKOLDPSLCTHLIYAFAGMTNHQLST-----TEW 71
Qy 74 -EKGYERFNNLRKNPELTMTISLGGWYEGSEKYSDMAANPTYRQFIQSVDLFLQYK 132
Db 72 NDETLYOEFNGLKKNPKLTLAIGWNFTQKFTDMVATANNRQTFVNSAIRLKYK 131
Qy 133 FDGLDLWEYPGSRIGNPKDKONYALVRELKDAFEPHG-----YLLTAASVPGDK 185
Db 132 FDGLDLWEYPGSQ--GSPAVDKERTTLVQDLANAFQOEAOQSGKERLLLSAAVPAGQY 190
Qy 186 IDRAYDIKELNKLFDWNVMTYDYGWENFYGHNAPLYKRPDDELTHTYFNVTMTHY 245
Db 191 VDAGYEVDKIAQNLDVNLVAYDFHGSWEKVTGHNSPLYKREQESGAAS-LNVDAAVQ 249
Qy 246 YLNGATRDKLVMPVPGYGRWSIEDRSKLGDPKAGMSPPGFISGEEGVLSYIELCQL 305
Db 250 WLQKGTTPASKLILGMPTYSFSLASSSDTRVGAPATGSGTGPFTKEGGLMAYEVCWS 309
Qy 306 FQKEEWHIOYDEYVYNAFYNDKIWGYDDLASISCKLAFKELGVSVMVMSLENDDEFK 365
Db 310 KGATKQRIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSKLQKGLGAMVWALDDFA 366
Qy 366 GHCGPKNPLLNKVNHMGDEKNSFECILGPGSTTTPTTPTTPTTPTTPTTPTTPTTPTT 425
Db 367 G-----FSCNQ-----RYPLIQTLRQELSPLPLPSGT 394

Qy 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450
Db 395 PELEVP-KPGQSPSEPHGSPGQDT 418

RESULT 10
US-10-004-219B-9
; Sequence 9, Application US/10004219B
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aeris, Johannes M.F.G.

```

; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(452)
; OTHER INFORMATION: /note="Mouse AMCase"
US-10-004-219B-9

Query Match      24.2%; Score 729.5; DB 14; Length 452;
Best Local Similarity 37.0%; Pred. No. 3.4e-45;
Matches 154; Conservative 69; Mismatches 144; Indels 49; Gaps 11;

QY      16  IVCYGVTSVYHK-VDPTTIEDIDPFKCTHLMYGFADKIDYKTIQVDPYQDDNHNW- 73
DB      3  LICYFTNWAQYRPGGLSGFKPDDINPCLCTHLIYAFAGQNNNEITT-----IEWN 51
QY      74  EKRGYERFNNLRKNPELTMTTMSLGSWYEGSEKYSDMAANPTYRQOFIQSVLDFLOEYKF 133
DB      52  DVTLYKAENLKNRNSKULTLAIAGWNFGTAPFTTWTSTQNRQTFTTSVKFLRQYGF 111
QY      134  DGLDLDWEYPGSRGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASVPGKDKI 186
DB      112  DGLDLDWEYPGSR- GSPPODKHLFTVLVKNREAFQEAEIESNRRLMTVAAGGINSI 170
QY      187  DRAYDIKELNKLFDMMNVNTYDYGWENFYGHNAPLYKRPDETDLHTYFNNVTMYY 246
DB      171  QAGYEIPELSKYLDPIHVMYDHLGSEWGYTGENSEPLYKYPTETGS-NAYLNVYVMYVW 229
QY      247  LNNGATRDKLVMGVFFYGRANSIEDRSKLGDPKAGNSPPGFISGEGVLSYIELCOLF 306
DB      230  KNGGAPAEKLVGPPEYGHFTILRNPSDNGIGAPTSGDPAGAYTROAGFWAYYEICTFL 289
QY      307  QK---EEWHIQDEYNAPYGVNDKIWGYDDLASISCKLAFKELGVSQVWVWSLEND 363
DB      290  RSGATEVW----DASQEVPIAYKANWLGYNIKSFSVKAQWLKONNFGGAMIWALDLD 345
QY      364  FKQ-HCG-----PKNPLLNKVNMMINGDEKNSFCILGPESTTTPTTPTTPTTTP 414
DB      346  FTGSPCDQKQKFLTSTLNKA-----LGISTEGCTAPDVPSEPVITP 386

RESULT 11
US-10-004-219B-4
; Sequence 4, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: amino acid sequence deduced from cDNA sequence
US-10-004-219B-4

Query Match      24.2%; Score 729.5; DB 14; Length 473;
Best Local Similarity 37.0%; Pred. No. 3.6e-45;
Matches 154; Conservative 69; Mismatches 144; Indels 49; Gaps 11;

QY      16  IVCYGVTSVYHK-VDPTTIEDIDPFKCTHLMYGFADKIDYKTIQVDPYQDDNHNW- 73
DB      24  LICYFTNWAQYRPGGLSGFKPDDINPCLCTHLIYAFAGQNNNEITT-----IEWN 72
QY      74  EKRGYERFNNLRKNPELTMTTMSLGSWYEGSEKYSDMAANPTYRQOFIQSVLDFLOEYKF 133
DB      73  DVTLYKAENLKNRNSKULTLAIAGWNFGTAPFTTWTSTQNRQTFTTSVKFLRQYGF 132
QY      134  DGLDLDWEYPGSRGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASVPGKDKI 186
DB      133  DGLDLDWEYPGSR- GSPPODKHLFTVLVKNREAFQEAEIESNRRLMTVAAGGINSI 191
QY      187  DRAYDIKELNKLFDMMNVNTYDYGWENFYGHNAPLYKRPDETDLHTYFNNVTMYY 246
DB      192  QAGYEIPELSKYLDPIHVMYDHLGSEWGYTGENSEPLYKYPTETGS-NAYLNVYVMYVW 250
QY      247  LNNGATRDKLVMGVFFYGRANSIEDRSKLGDPKAGNSPPGFISGEGVLSYIELCOLF 306
DB      251  KNGGAPAEKLVGPPEYGHFTILRNPSDNGIGAPTSGDPAGAYTROAGFWAYYEICTFL 310
QY      307  QK---EEWHIQDEYNAPYGVNDKIWGYDDLASISCKLAFKELGVSQVWVWSLEND 363
DB      311  RSGATEVW----DASQEVPIAYKANWLGYNIKSFSVKAQWLKONNFGGAMIWALDLD 366
QY      364  FKQ-HCG-----PKNPLLNKVNMMINGDEKNSFCILGPESTTTPTTPTTPTTTP 414
DB      367  FTGSPCDQKQKFLTSTLNKA-----LGISTEGCTAPDVPSEPVITP 407

RESULT 12
US-10-161-547-14
; Sequence 14, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-14

Query Match      23.9%; Score 721.5; DB 14; Length 373;
Best Local Similarity 37.7%; Pred. No. 1e-44;
Matches 136; Conservative 83; Mismatches 117; Indels 25; Gaps 8;

QY      15  RIVCYGVTSVYHKVDP-YTIEDIDPFKCTHLMYGFADKIDYKTIQVDPYQDDNHNW 73
DB      2  KLVICYFTNWAQYRQGEARFLPKDLPDLCTHLIYAFAGTNNHQLST-----TEW 50
QY      74  -EKRGYERFNNLRKNPELTMTTMSLGSWYEGSEKYSDMAANPTYRQOFIQSVLDFLOEYKF 132
DB      51  NDETLYQEENGLKKNPKLKTLLAIGWNFGTQKFTDMVATANNRQTFTVNSAIRFLKYS 110
QY      133  FDGLDLDWEYPGSRGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAASVPGKDKI 185

```

```
Db 111 FDGLDLDWEYFGSQ--GSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTY 169
QY 186 IDRAYDIKELNKLFDWNNVNTYDYGWENFYGHNAFLYKRPDETDELHYFNNTMYH 245
Db 170 VDAGEYVDKIAQNLDLFVNLAYDFHGSWEKVTGHSPLRYKQESGAAS--LNVDAAVQ 228
QY 246 YLNGATRDKLVMGVPYFGRAWSEDRSKLGLDPAKGMSPPGFISGEEGVLSYIELCOL 305
Db 229 WLQKGTASKLILGMPTYGHSFTLASSSDTRVGAPATGSGTGPFTFKEGMLAYEVCWS 288
QY 306 FQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVLGVGMVWVSLNDDEPK 365
Db 289 KGATKQRIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDDEFA 345
QY 366 G 366
Db 346 G 346

RESULT 13
US-10-161-547-15
; Sequence 15, Application US/10161547
; Publication No. US20030143216A1
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/10/161,547
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/267,574
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/039,198
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-547-15

Query Match 23.9%; Score 721.5; DB 14; Length 373;
Best Local Similarity 37.7%; Pred. No. 1e-44;
Matches 136; Conservative 83; Mismatches 117; Indels 25; Gaps 8;

QY 15 RIVCYVGTWSVYHKVDP-YIIEDPDKCTHLMYGFAKIDYKTYIQVDFPYQDDNHSW 73
Db 2 KLVCYFTNAQYRQGEARFLPKDLPSCLTHLIYAFAGMTNHLST-----TEW 50
QY 74 -EKGYERENLRNLKNPELTMTISLGGWYEGSEKYSMDAANPTYRQOFIQSVLDFLOEYK 132
Db 51 NDETLYOEFNGLKKNPKLTLAIGWNGFTQKFTDMVATANNROTFTVNSAIRFLUKYS 110
QY 133 FDGLDLDWEYFGSRLGNPKDKQNYLALVRELKDAPEPHG-----YLLTAAVSPGKDK 185
Db 111 FDGLDLDWEYFGSQ--GSPAVDKERFTTLVQDLANAFQEAQTSKGERLLLSAAVPAGQTY 169
QY 186 IDRAYDIKELNKLFDWNNVNTYDYGWENFYGHNAFLYKRPDETDELHYFNNTMYH 245
Db 170 VDAGEYVDKIAQNLDLFVNLAYDFHGSWEKVTGHSPLRYKQESGAAS--LNVDAAVQ 228
QY 246 YLNGATRDKLVMGVPYFGRAWSEDRSKLGLDPAKGMSPPGFISGEEGVLSYIELCOL 305
Db 229 WLQKGTASKLILGMPTYGHSFTLASSSDTRVGAPATGSGTGPFTFKEGMLAYEVCWS 288
QY 306 FQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVLGVGMVWVSLNDDEPK 365
Db 289 KGATKQRIQ-DQ--KVPYIFRDNQWGFDDVESFKTKVSYLKQKGLGAMVWALDLDDEFA 345
QY 366 G 366
```

```
Db 346 G 346

RESULT 14
US-10-004-219B-14
; Sequence 14, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aeris, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(455)
; OTHER INFORMATION: /note="Human AMCase"
US-10-004-219B-14

Query Match 23.5%; Score 708.5; DB 14; Length 455;
Best Local Similarity 35.9%; Pred. No. 1.2e-43;
Matches 149; Conservative 75; Mismatches 154; Indels 37; Gaps 10;

QY 15 RIVCYVGTWSVYHK-VDPYTTIEDIPFKCTHLMYGFAKIDYKTYIQVDFPYQDDNHSW 73
Db 2 QLTCTFTWAQYRPGLGFMFDNIDPCLCTHLIYAFAGRONNEITT-----IEW 50
QY 74 -EKGYERENLRNLKNPELTMTISLGGWYEGSEKYSMDAANPTYRQOFIQSVLDFLOEYK 132
Db 51 NDETLYOEFNGLKKNPKLTLAIGWNGFTQKFTDMVATANNROTFTVNSAIRFLUKYS 110
QY 133 FDGLDLDWEYFGSRLGNPKDKQNYLALVRELKDAPEPHG-----YLLTAAVSPGKDK 185
Db 111 FDGLDLDWEYFGSR--GSPPDQKHLFTVLVQEMREAFQEAQINKRPLMTAAVAAGISN 169
QY 186 IDRAYDIKELNKLFDWNNVNTYDYGWENFYGHNAFLYKRPDETDELHYFNNTMYH 245
Db 170 IQSGYEIPQLSOYLDYIHMTYDLHGSWEGYTGNSPLYKYPTDTGS--NAVLYNDVYVNY 228
QY 246 YLNGATRDKLVMGVPYFGRAWSEDRSKLGLDPAKGMSPPGFISGEEGVLSYIELCOL 305
Db 229 WKDNGAPAEKLVGFTYGHNFILSNPNTGIGFTSAGAPGVAKESGWAYVEIC-T 287
QY 306 FQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVLGVGMVWVSLNDDEPK 365
Db 288 FLKNGATQGDAPQEPVPYAYQGNVWVGYDNKSPDKAQLKHNKFGAMVWALDLDFT 347
QY 366 G-HCGP-KNPLKNVHNNINGDEKNSFECILGPSITTTPTPTPTPTPTPTPTPTPTPS 418
Db 348 GTFCNQCKFPPLT-----STLKKALGLQASCTAPAPQETPETAAPSGS 390

RESULT 15
US-10-004-219B-1
; Sequence 1, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aeris, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: which mucus is involved or infection diseases
```

Search completed: March 22, 2004, 07:45:53
Job time : 110.643 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:30:23 ; Search time 162.975 Seconds
(without alignments)
929.256 Million cell updates/sec

Title: US-09-662-293-21

Perfect score: 3014

Sequence: 1 SIKEDHNDYKNPNIYCV.....IMPCPPTQWQKLTQIGE 536

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|---------------------|
| 1 | 3014 | 100.0 | 536 | 3 AAY52525 | Aay52525 House dus |
| 2 | 3014 | 100.0 | 536 | 5 AAU96329 | Aau96329 Der HW-m |
| 3 | 3014 | 100.0 | 555 | 3 AAY52523 | Aay52523 House dus |
| 4 | 3014 | 100.0 | 555 | 5 AAU96327 | Aau96327 Der HW-m |
| 5 | 3014 | 100.0 | 555 | 5 AAU96328 | Aau96328 Der HW-m |
| 6 | 2475 | 82.1 | 490 | 3 AAY52535 | Aay52535 D. pteron |
| 7 | 2475 | 82.1 | 490 | 5 AAU96339 | Aau96339 Der HW-m |
| 8 | 2475 | 82.1 | 509 | 3 AAY52533 | Aay52533 D. pteron |
| 9 | 2475 | 82.1 | 509 | 5 AAU96337 | Aau96337 Der HW-m |
| 10 | 2475 | 82.1 | 509 | 5 AAU96338 | Aau96338 Der HW-m |
| 11 | 1007 | 33.4 | 525 | 6 ABP72636 | Abp72636 Anopheles |
| 12 | 848.5 | 28.2 | 4498 | 4 ABB58595 | Abb58595 Drosophil |
| 13 | 837.5 | 27.8 | 467 | 6 ABP72634 | Abp72634 Prawn chit |
| 14 | 837.5 | 27.8 | 595 | 4 ABB71737 | Abb71737 Drosophil |
| 15 | 834 | 27.7 | 554 | 2 AAU01824 | Aau01824 Manduca s |
| 16 | 834 | 27.7 | 554 | 3 AAB07183 | Aab07183 Manduca s |
| 17 | 834 | 27.7 | 554 | 6 ABP72619 | Abp72619 Manduca s |
| 18 | 830.5 | 27.6 | 460 | 4 ABB64366 | Abb64366 Drosophil |
| 19 | 819 | 27.2 | 565 | 6 ABP72625 | Abp72625 Bombyx mo |
| 20 | 812.5 | 27.0 | 583 | 5 AAE28197 | Aae28197 Flea chit |
| 21 | 811.5 | 26.9 | 559 | 5 AAE28199 | Aae28199 Flea chit |
| 22 | 809.5 | 26.9 | 635 | 5 AAE28203 | Aae28203 Flea chit |
| 23 | 808 | 26.8 | 574 | 6 ABP72635 | Abp72635 Aedes aeg |
| 24 | 801.5 | 26.6 | 483 | 6 ABP72633 | Abp72633 Chelonus |
| 25 | 779.5 | 25.9 | 553 | 6 ABP72626 | Abp72626 Hyphantri |

| | | | | | |
|----|-------|------|-----|------------|--------------------|
| 26 | 738 | 24.5 | 466 | 2 AAW08584 | Aaw08584 Human 50 |
| 27 | 738 | 24.5 | 466 | 2 AAW40259 | Aaw40259 Human chi |
| 28 | 738 | 24.5 | 466 | 2 AAY42425 | Aay42425 MO-218 cl |
| 29 | 738 | 24.5 | 466 | 4 AAE00432 | Aae00432 Human chi |
| 30 | 738 | 24.5 | 466 | 5 AAE25903 | Aae25903 Human chi |
| 31 | 738 | 24.5 | 466 | 5 ABB76291 | Abb76291 Human chi |
| 32 | 735.5 | 24.4 | 459 | 7 ADC24231 | Adc24231 Human NOV |
| 33 | 732.5 | 24.3 | 387 | 2 AAW08585 | Aaw08585 Human 39 |
| 34 | 732 | 24.3 | 447 | 7 ADC24237 | Adc24237 Human NOV |
| 35 | 732 | 24.3 | 466 | 2 AAW40260 | Aaw40260 Human NOV |
| 36 | 732 | 24.3 | 466 | 2 AAY42426 | Aay42426 MO-13B cl |
| 37 | 732 | 24.3 | 466 | 4 AAE00433 | Aae00433 Human chi |
| 38 | 732 | 24.3 | 466 | 5 AAE25904 | Aae25904 Human chi |
| 39 | 732 | 24.3 | 466 | 5 ABB76292 | Abb76292 Human chi |
| 40 | 731 | 24.3 | 466 | 2 AAW31498 | Aaw31498 Human chi |
| 41 | 729.5 | 24.2 | 452 | 7 ABR55544 | Abr55544 Amino aci |
| 42 | 729.5 | 24.2 | 473 | 7 ABR55543 | Abr55543 Amino aci |
| 43 | 729.5 | 24.2 | 520 | 6 ABU09914 | Abu09914 Partial m |
| 44 | 727.5 | 24.1 | 473 | 7 ADC51464 | Adc51464 Chitotria |
| 45 | 723 | 24.0 | 466 | 6 ABP72621 | Abp72621 Human chi |

ALIGNMENTS

RESULT 1

AAV52525

ID AAY52525 standard; protein; 536 AA.

XX AAY52525;

XX 22-FEB-2000 (first entry)

XX House dust mite (D. farinae) mite allergen protein (map) PDerf98-536.

XX Mite allergen protein; map; high molecular weight; HMW-map; allergy;

XX house dust mite; IgE; immunoglobulin E; allergen; mapA; mapB;

XX hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;

XX canine; veterinary; antibody; vaccine; immunisation.

XX Dermatophagoides farinae.

XX WO954349-A2.

XX 28-OCT-1999.

XX 16-APR-1999; 99WO-US008524.

XX 17-APR-1998; 98US-00062013.

XX 13-MAY-1998; 98US-0085293P.

XX 02-SEP-1998; 98US-0098909P.

XX (HESK-) HESKA CORP.

XX McCall CA, Hunter SW, Weber ER;

XX WPI; 2000-052700/04.

XX N-PSDB; AA238579, AA238580.

XX Novel high molecular weight Dermatophagoides nucleic acid polypeptides

XX used to modify an animals' hypersensitivity to mite allergens.

XX Claim 3; Page 125-127; 154pp; English.

XX This sequence represents Dermatophagoides farinae mite allergen protein

XX (map) PDerf98-536, the mature form of PDerf98-555 (AAY52523). PDerf98-536

XX has a molecular weight of 98 kD, comprising 536 amino acids, and is a

XX component of the Dermatophagoides farinae high molecular weight mite

XX allergen protein (HMW-map) composition. The HMW-map composition was

XX isolated from a D. farinae homogenate by gel filtration, with each

XX fraction being analysed for the presence of proteins that bound to IgE

XX present in mite-allergic dog antisera. Mite allergenic proteins and

XX peptides, and nucleic acids encoding them, may be used in therapeutic

CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
 CC acid. The Der HMW-map protein is useful for eliciting an immune response
 CC against Der HMW-map protein. The protein or a reagent comprising a non-
 CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
 CC cat) susceptible to or having an allergic response to a mite. A
 CC therapeutic composition is useful for desensitising a host animal to an
 CC allergic response to a mite. The DNA and protein can be used in the
 CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
 CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
 CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
 CC binding of proteins to IgE, to prevent immunocomplex formation, thus
 CC reducing hypersensitivity responses to mite allergens, and as vaccines
 CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
 CC represent Der HMW-map polypeptides of the invention
 XX
 XX SQ Sequence 555 AA;

Query Match 100.0%; Score 3014; DB 5; Length 555;
 Best Local Similarity 100.0%; Pred. No. 3.9e-210;
 Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIKRDNDYKSNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 60
 Db 20 SIKRDNDYKSNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 79
 QY 61 VFDPYQDDNHNHSEKRGYERFNNLKNPELTMTISLGGWTEGSEKYSDMAANPTYRQF 120
 Db 80 VFDPYQDDNHNHSEKRGYERFNNLKNPELTMTISLGGWTEGSEKYSDMAANPTYRQF 139
 QY 121 IQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAS 180
 Db 140 IQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAS 199
 QY 181 PGKOKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 240
 Db 200 PGKOKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 259
 QY 241 YTMHYLLNNGATRDKLVMGVPPFYGRAMSIEDRSKILKLGDPKAGMSPPGFISGEEGLSYI 300
 Db 260 YTMHYLLNNGATRDKLVMGVPPFYGRAMSIEDRSKILKLGDPKAGMSPPGFISGEEGLSYI 319
 QY 301 ELCOLFQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVSGVMVWSLE 360
 Db 320 ELCOLFQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVSGVMVWSLE 379
 QY 361 NDDFKGCHGCPKPNLLKNVHNMINGEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTT 420
 Db 380 NDDFKGCHGCPKPNLLKNVHNMINGEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTT 439
 QY 421 TPTT 480
 Db 440 TPTT 499
 QY 481 VDGLIKYKCGDIPHPNTIHKYLVCFVNGVWVHMPCPGPIWQCKLTICGE 536
 Db 500 VDGLIKYKCGDIPHPNTIHKYLVCFVNGVWVHMPCPGPIWQCKLTICGE 555

RESULT 5
 AAU96328
 ID AAU96328 standard; protein; 555 AA.
 XX
 XX AC AAU96328;
 XX
 XX DT 15-JUL-2002 (first entry)
 XX
 XX DE Der HMW-map polypeptide #15.
 XX
 XX KW Der HMW-map; American house dust mite; antiallergic; mite; IgE;
 XX KW mite allergenic protein; immunoglobulin E; hypersensitivity;
 XX KW immunocomplex formation.
 XX
 XX OS Dermatophagoides farinae.

XX WO200222807-A2.
 XX 21-MAR-2002.
 XX 14-SEP-2001; 2001WO-US028730.
 XX 14-SEP-2000; 2000US-00662293.
 XX (HESK-) HESKA CORP.
 XX McCall CA, Hunter SW, Weber ER;
 XX WPI; 2002-351888/38.
 XX N-PSDB; ABK69573.
 XX
 XX PT New mite allergenic protein isolated from Dermatophagoides, designated
 XX Der HMW-map protein, useful as a vaccine for treating mite allergy.
 XX
 XX PS Claim 12; Page 120-122; 161pp; English.
 XX
 XX The invention relates to an isolated mite allergenic protein of
 CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
 CC acid. The Der HMW-map protein is useful for eliciting an immune response
 CC against Der HMW-map protein. The protein or a reagent comprising a non-
 CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
 CC cat) susceptible to or having an allergic response to a mite. A
 CC therapeutic composition is useful for desensitising a host animal to an
 CC allergic response to a mite. The DNA and protein can be used in the
 CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
 CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
 CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
 CC binding of proteins to IgE, to prevent immunocomplex formation, thus
 CC reducing hypersensitivity responses to mite allergens, and as vaccines
 CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
 CC represent Der HMW-map polypeptides of the invention
 XX
 XX SQ Sequence 555 AA;

Query Match 100.0%; Score 3014; DB 5; Length 555;
 Best Local Similarity 100.0%; Pred. No. 3.9e-210;
 Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIKRDNDYKSNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 60
 Db 20 SIKRDNDYKSNPMRIVCVGTWVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 79
 QY 61 VFDPYQDDNHNHSEKRGYERFNNLKNPELTMTISLGGWTEGSEKYSDMAANPTYRQF 120
 Db 80 VFDPYQDDNHNHSEKRGYERFNNLKNPELTMTISLGGWTEGSEKYSDMAANPTYRQF 139
 QY 121 IQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAS 180
 Db 140 IQSVLDFLOEYKFDGLDWEYPSGRLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAS 199
 QY 181 PGKOKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 240
 Db 200 PGKOKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 259
 QY 241 YTMHYLLNNGATRDKLVMGVPPFYGRAMSIEDRSKILKLGDPKAGMSPPGFISGEEGLSYI 300
 Db 260 YTMHYLLNNGATRDKLVMGVPPFYGRAMSIEDRSKILKLGDPKAGMSPPGFISGEEGLSYI 319
 QY 301 ELCOLFQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVSGVMVWSLE 360
 Db 320 ELCOLFQKEEMHIQYDEYNNAPYGYNDKIWGYDDLASISCKLAFKELGVSGVMVWSLE 379
 QY 361 NDDFKGCHGCPKPNLLKNVHNMINGEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTT 420
 Db 380 NDDFKGCHGCPKPNLLKNVHNMINGEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTT 439
 QY 421 TPTT 480

PT New mite allergenic protein isolated from Dermatophagoides, designated
 Der HWM-map protein, useful as a vaccine for treating mite allergy.
 XX
 XX
 XX Claim 12; Page 144-146; 161pp; English.
 CC
 CC The invention relates to an isolated mite allergenic protein of
 CC Dermatophagoides, designated Der HWM-map protein, and its related nucleic
 CC acid. The Der HWM-map protein is useful for eliciting an immune response
 CC against Der HWM-map protein. The protein or a reagent comprising a non-
 CC proteinaceous epitope is useful for identifying an animal (e.g., dog,
 CC cat) susceptible to or having an allergic response to a mite. A
 CC therapeutic composition is useful for desensitising a host animal to an
 CC allergic response to a mite. The DNA and protein can be used in the
 CC detection of anti-Der HWM-map antibodies in animal fluids, and inhibition
 CC of Immunoglobulin (Ig)E or Der HWM-map protein activity associated with a
 CC disease. Antibodies that bind to Der HWM-map are useful for inhibiting
 CC binding of proteins to IgE, to prevent immunocomplex formation, thus
 CC reducing hypersensitivity responses to mite allergens, and as vaccines
 CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
 CC represent Der HWM-map polypeptides of the invention

Sequence 490 AA;

| | | | | |
|-----------------------|--------------|---------------------|----------------|-------------|
| Query Match | 82.1%; | Score 2475; | DB 5; | Length 490; |
| Best Local Similarity | 82.5%; | Pred. No. 4.2e-171; | | |
| Matches 443; | Conservative | 21; | Mismatches 21; | Indels 52; |
| Gaps | 2; | | | |

[illegible]

QY 483 GHLIKCYREGDI PPTNIHKYLVEFV--NGGWWHIMPCPGTIWCQKLTIGE 536
||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Dd 434 GHLIKCYQGYLEPPTDVKLYLCEYIATNGGWWHIIMDCPKGRTHWATLKNCIOE 490

RESULT 8

AA52533

ID AAY52533 standard; protein; 509 AA.

AC AAY52533;

DT 06-AUG-2003 (revised)

DT 22-FEB-2000 (first entry)

X

| | |
|----|--|
| DE | D. pteronyssinus 98 kD mite allergen protein (map) PDerp98-509. |
| XX | |
| KW | Mite allergen protein; map; high molecular weight; HMW-map; allergy; |
| KW | house dust mite; IGE; immunoglobulin E; allergen; mapB; |
| KW | hypersensitivity reaction; therapy; treatment; diagnosis; human; feline; |
| KW | canine; veterinary; antibody; vaccine; immunisation. |
| XX | |
| OS | Dermatophagoideis pteronyssinus. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | Peptide 1..19 |
| FT | /note= "Signal peptide" |
| FT | Protein 20..509 |
| FT | /note= "Mature PDerp98-509" |
| XX | |
| PN | WO9954349-A2. |
| XX | |
| PD | 28-OCT-1999. |
| XX | |
| PF | 16-APR-1999; 99WO-US008524. |
| XX | |
| PR | 17-APR-1998; 98US-00062013. |
| PR | 13-MAY-1998; 98US-0085295P. |
| PR | 02-SEP-1998; 98US-0098909P. |
| XX | |
| PA | (HESK-) HESKA CORP. |
| XX | |
| PI | Mccall CA, Hunter SW, Weber ER; |
| XX | |
| WI | WPI; 2000-052700/04. |
| DR | N-PSDB; AAZ38585, AAZ38586, AAZ38587, AAZ38588. |
| XX | |
| PT | Novel high molecular weight Dermatophagoides nucleic acid polypeptides |
| PT | used to modify an animals' hypersensitivity to mite allergens. |
| XX | |
| PS | Claim 3; Page 134-136; 154pp; English. |

PS Claim 3; Page 134-136; 154pp; English.

This sequence represents Dermatophagoides pteronyssius mite allergen protein (map). Pterp99-509 has a molecular weight of 98 kD, comprising 509 amino acids, and has a high degree of homology with the D. farinae 98 kD allergen, mapB (AAV52523). Nucleic acid molecules encoding Pterp99-509 were isolated from a D. pteronyssius cDNA library by hybridisation with a probe encoding the D. farinae high molecular weight map (Hmw-map) composition. Mite allergenic proteins and peptides, and nucleic acids encoding them, may be used in therapeutic compositions to modify an animal's hypersensitivity reaction to mite allergens. Animals that may be treated include mammals and birds, especially felines, canines, equines, humans, other pets, and work or domestic animals. The proteins or fragments may also be used to diagnose allergies via a skin test. The proteins and peptides can also be used to raise antibodies, which have a variety of potential uses. For example, they can be used as vaccines to passively immunise animals against dust mite hypersensitivity, as positive controls in test kits and as tools to recover desired dust mite allergens from a mixture of proteins. (Updated on 06-AUG-2003 to correct OS field.)

Sequence 509 AA:

Query Match 82.1%; Score 2475; DB 3; Length 509;

Best Local Similarity 82.5%; Pred. No. 4.4e-171;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
Length 309;

QY 3 KRDN DY SKN PMR IVCYVGTWSVYHKVDPYTIEDIDPEKCTHLMYGF AKIDV KYTIOVF 62

Db 22 KRDNHNSKNPMRIVCVGTWSVYHKVDPTYIEDIDPFKCTHLMYGFAKIDEXYKTIOVF 81

QY 63 DPYQDDNHSWEKRGYERFNNLRLKNPELT'TTMSLGGWYEGSEKYSMDMAANPTYROOFIO 122

Db 82 DPQDDNNHNSKCHGYERFNNRLKNPELTTWISLGGWYEGSEKYSIDMAANPTVROQFVQ 141

QY 123 SVLDFLOEYKFDGLDLDWEYFGSRLGNPKIDKONYLALVRELKDAFEPHGYLLTAAVSPG 182

Db 142 SVLDFLOEYKEDGLDLDWEYPSGRLGNPKDKONYLTTLVRELKEAFEPFGYLTAAVSPG 201

Db 82 YLDLEEN-WGEGHAKRFGVGLKNVGFGLKTLAAIGWNEGRSKSAMASGELKRRFISDC 140
Qy 125 LDFQYKFGDGLDWEYPSRSLGNPKIDKONYALVRELKDAPEHGYLLTAAVSPGKD 184
Db 141 VAFQQRHGFQDGLDWEYPAQRDGNPLIDRDNAQLVEEMREBFDHYGLLLTAAVASVEF 200
Qy 185 KIDRAYDIKELNKLFDWNNVTYDYGWENFYGHNAPLYKRPDETDDELHTYFNVNYTMH 244
Db 201 SAGVSYDIPRISKSFHFLNVMVDMHGAWDSYOCINAPLYRGSAADTDRLGQINVASIH 260
Qy 245 YLLNNGATRDKLVMGVFPYGRANSIEDRSKLKGDPAKMSPPGFTSGEGVLSYELCQ 304
Db 261 FWLAQGGCTGRKLVIGLIFLYGRNFTLASAANTQIGAPTGGGTGRTREFGVNGYNEFCE 320
Qy 305 LFOKEEWHIOYDEYNAAPYGVNDKIWGYDDLASISCKLAFKLKELGVGMWWSLNDDF 364
Db 321 KLATEAWDLWSEQQVPIAVRNQWGVYDLSVQLKVYLDDQGLGGAMWSLSTDF 380
Qy 365 KGHG-GPKNPLLNKVNMMINGDEKNSFECILGFTSTTPTTPTTPTTPTTPTTPTTPTT 423
Db 381 LGVCGGGRYPLMHEIRSLVNGGT-----FSTTTPPSVAPT-----STVAGTTT 426
Qy 424 TTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 483
Db 427 TTPGANPGTQPTT--SDAPNHTTSTTTEGNGFTTRPSG-----DQ 468
Qy 484 HLKCY--KEGDIHPNINHKLVC-----EFVNGWVHIMPCPGTIW 526
Db 469 ---PCAGRGYGFVPHPTNCARYICLTADTYEFT-----CPPGTLP 507

RESULT 12

ABBS8595
ID ABBS8595 standard; protein; 4498 AA.
XX AC ABBS8595;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 2577.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-656860/75.
DR N-PSDB; ABL02698.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 2577; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS5773-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4498 AA;
Query Match 28.2%; Score 848.5; DB 4; Length 4498;
Best Local Similarity 34.9%; Pred. No. 7.3e-52;
Matches 197; Conservative 81; Mismatches 18; Indels 107; Gaps 19;
Qy 15 RIVCVGVWTVVH-----KVDPTIIDIPEKCTHLMY---GPAKIDVYKTIQVDPYQD 67
Db 53 RVVCYTWSVYRPTAKFNP---QNIPIYCLTHLVYAFGGFTKDNQMK---PDKYQD 105
Qy 68 DNHNSWEKRGYERFNNLRKLNPELTMTISLQGVTEGSEKYSDMAANPYRQOFTQSVLDF 127
Db 106 -----IEGGYAKFTGLATYKQLKTMIAIGWNEASRFPPLVASNERRQOFTKILKF 160
Qy 128 LOEYKFDGLDWEYPSRSLGNPKIDKONYALVRELKDAPEHGH-----YLLTAVS 180
Db 161 LEQNHFDGIDLWEYPAHREGGKSRDRDNYAQFVQELRAEFERAEKTRTLLTMAVP 220
Qy 181 PKDKIDRAYDIKELNKLFDWNNVTYDYGWENFYGHNAPLYKRPDETDDELHTY---F 237
Db 221 AGIEYIDKGYDVPKLNKYLDWFLVLYDFHSHSHEPSVNHAPLYSL--EEDSEYNYDAEL 278
Qy 238 NVNVTMYLLNNGATRDKLVMGVFPYGRANSIEDRSKLKGDPAKMSPPGFTSGEGVL 297
Db 279 NIDYSIKYLLXAGADRDKLVIGIPYGRSYTLNEESTELGAPAEQGEQDAREKGYL 338
Qy 298 SYIELCQLFOKE-EWH-IQYDEYNAAPYGVNDKIWGYDDLASISCKLAFKLKELGVSGVM 355
Db 339 AYIELCQLKDDPEWTVVQPNANVMGYAYRNQWGVYDDDEAIYKKAAYVVAQGLGIM 398
Qy 356 VWSLENDDFKHC-GPKNPLLNKVNMMINGDEKNSFECILGFTSTTPTTPTTPTTPTTPTT 414
Db 399 FWAINDDDFRGTGCKGPKYPLTEAA-----KEAMVEALGLGINEVAKPSGPKPSRSR 450
Qy 415 T-----TPSFTT---PTTTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 443
Db 451 SRDNASNRNLNGKTEAPLSRRRSATRRPAVSSTQAPP--PSTFKLTEAGSSLYTGG 508
Qy 444 -PSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 502
Db 509 RASTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 542
Qy 503 YLVCFEYNG-----GWWVHIMPCPPG 523
Db 543 YYWC-LDSGPGSLGIVAHMFTCPG 566

RESULT 13

ABP72634
ID ABP72634 standard; protein; 467 AA.
XX AC ABP72634;
XX DT 23-OCT-2003 (revised)
XX DT 11-JUN-2003 (first entry)
XX DE Prawn chitinase.
XX KW Insecticide; pesticide; insect control; insect; toxin; chitinase; enzyme;
XX KW neuropeptide; transgenic plant; crop protection; prawn.
XX OS Marsupenaeus japonicus.
XX FH Key Location/Qualifiers
XX FT Active-site 118..125

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:39:53 ; Search time 107.855 Seconds
(without alignments)
1568.003 Million cell updates/sec

Title: US-09-662-293-21
Perfect score: 3014
Sequence: 1 SIKRDNDYSKNPBRIVCV.....IMPCPPGTWQEKLTICGE 536

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviris:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 3014 | 100.0 | 555 | Q9U6R7 | Q9U6R7 dermatophag |
| 2 | 1007 | 33.4 | 525 | O44079 | O44079 anopheles g |
| 3 | 938 | 31.1 | 431 | Q81SH5 | Q81SH5 araneus ven |
| 4 | 869 | 28.8 | 1635 | O17412 | O17412 aedes aegyp |
| 5 | 848.5 | 28.2 | 4498 | Q9W2Z3 | Q9W2Z3 drosophila |
| 6 | 842 | 27.9 | 2838 | Q8MP05 | Q8MP05 tenebrio mo |
| 7 | 837.5 | 27.8 | 467 | O15993 | O15993 penaeus jap |
| 8 | 837.5 | 27.8 | 595 | Q9VFR3 | Q9VFR3 drosophila |
| 9 | 831.5 | 27.6 | 467 | Q81TU3 | Q81TU3 penaeus van |
| 10 | 830.5 | 27.6 | 460 | Q9W2M7 | Q9W2M7 drosophila |
| 11 | 826 | 27.4 | 544 | Q9GQC4 | Q9GQC4 bombyx mori |
| 12 | 823 | 27.3 | 488 | Q30W34 | Q30W34 bufo japoni |
| 13 | 821.5 | 27.3 | 557 | Q8MTK0 | Q8MTK0 choriostoneu |
| 14 | 819 | 27.2 | 543 | Q9GR93 | Q9GR93 bombyx mori |
| 15 | 819 | 27.2 | 543 | Q9GV05 | Q9GV05 bombyx mori |
| 16 | 819 | 27.2 | 565 | P90710 | P90710 bombyx mori |

| | | | | | | |
|----|-------|------|------|----|--------|--------------------|
| 17 | 818 | 27.1 | 566 | 5 | Q8WR52 | Q8WR52 bombyx mori |
| 18 | 811 | 26.9 | 565 | 5 | Q9GPG9 | Q9GPG9 bombyx mand |
| 19 | 808 | 26.8 | 574 | 5 | O17411 | O17411 aedes aegyp |
| 20 | 801.5 | 26.6 | 483 | 5 | Q23737 | Q23737 chelonus sp |
| 21 | 801 | 26.6 | 552 | 5 | Q9GV44 | Q9GV44 spodoptera |
| 22 | 797 | 26.4 | 572 | 5 | Q26042 | Q26042 penaeus jap |
| 23 | 790 | 26.2 | 620 | 5 | Q9Y0D4 | Q9Y0D4 penaeus mon |
| 24 | 779.5 | 25.9 | 553 | 5 | P91731 | P91731 hyphantria |
| 25 | 775 | 25.7 | 470 | 13 | Q803B7 | Q803B7 brachydanio |
| 26 | 755.5 | 25.1 | 474 | 5 | Q86L22 | Q86L22 lutzomyia l |
| 27 | 754 | 25.0 | 500 | 13 | Q7ZV48 | Q7ZV48 brachydanio |
| 28 | 738 | 24.5 | 466 | 4 | Q13231 | Q13231 homo sapien |
| 29 | 732.5 | 24.3 | 387 | 4 | Q9H3V8 | Q9H3V8 homo sapien |
| 30 | 732 | 24.3 | 482 | 13 | Q8AV87 | Q8AV87 gallus gall |
| 31 | 730.5 | 24.2 | 472 | 6 | Q95M17 | Q95M17 bos taurus |
| 32 | 729.5 | 24.2 | 473 | 11 | Q99PH2 | Q99PH2 mus musculu |
| 33 | 727.5 | 24.1 | 472 | 11 | Q9JLN1 | Q9JLN1 mus musculu |
| 34 | 727.5 | 24.1 | 473 | 11 | Q9D803 | Q9D803 mus musculu |
| 35 | 723.5 | 24.0 | 688 | 5 | Q8MS85 | Q8MS85 drosophila |
| 36 | 722.5 | 24.0 | 1013 | 5 | Q960M0 | Q960M0 drosophila |
| 37 | 717 | 23.8 | 527 | 5 | P91773 | P91773 penaeus jap |
| 38 | 711 | 23.6 | 460 | 5 | Q8WS95 | Q8WS95 glossina mo |
| 39 | 710 | 23.6 | 484 | 5 | Q9W092 | Q9W092 drosophila |
| 40 | 708.5 | 23.5 | 476 | 4 | Q9BZP6 | Q9BZP6 homo sapien |
| 41 | 702.5 | 23.3 | 929 | 5 | Q8MY79 | Q8MY79 haemaphysal |
| 42 | 702.5 | 23.3 | 983 | 5 | Q9VZV2 | Q9VZV2 drosophila |
| 43 | 687 | 22.8 | 381 | 11 | Q99J84 | Q99J84 mus musculu |
| 44 | 687 | 22.8 | 389 | 11 | Q8BKL8 | Q8BKL8 mus musculu |
| 45 | 677 | 22.5 | 461 | 5 | O01936 | O01936 chironomus |

ALIGNMENTS

RESULT 1

Q9U6R7
ID Q9U6R7 PRELIMINARY; PRT; 555 AA.
AC Q9U6R7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 98kDa HDM allergen.
OS Dermatophagoides farinae (House-dust mite).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;
OC Pyroglyphidae; Dermatophagoides.
OX NCBI_TaxID=6954;
RN [1]
RP SEQUENCE FROM N.A.
RA Weber B.R., Hunter S., Stedman K., McCall C.;
RT "Cloning and Characterization of a 98 kDa Allergen from
RT Dermatophagoides farinae";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF178772; AAD52672.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind PerA.
DR InterPro; IPR001223; Glyco Hydro 18.
DR InterPro; IPR001579; Glyco hydro 18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitinB2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 555 AA; 63238 MW; 0E4564A1A459B30B CRC64;

Query Match 100.0%; Score 3014; DB 5; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.2e-180;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SIKRDNDYKNPMEIVCVGTWTSVYHKVDPTTIEDIDPFKCTHLMYGFADIKDYKTYIQ 60
DB 20 SIKRDNDYKNPMEIVCVGTWTSVYHKVDPTTIEDIDPFKCTHLMYGFADIKDYKTYIQ 79
QY 61 VDPYODDNNHNSWEKRGYERFNNLRLKNPELTMTISLGGWYEGSEKYSMDAANTPTYRQOF 120
DB 80 VDPYODDNNHNSWEKRGYERFNNLRLKNPELTMTISLGGWYEGSEKYSMDAANTPTYRQOF 139
QY 121 IQSVLDFPQYKFDGLDLDWEYFSGSRGNPKIDKQNYLALVRELKDAPEPHGYLLTAAS 180
DB 140 IQSVLDFPQYKFDGLDLDWEYFSGSRGNPKIDKQNYLALVRELKDAPEPHGYLLTAAS 199
QY 181 PKQKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 240
DB 200 PKQKIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVN 259
QY 241 YTMHYLLNNGATRDKLVMGVPPFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEGVLSYI 300
DB 260 YTMHYLLNNGATRDKLVMGVPPFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEGVLSYI 319
QY 301 ELQCFQKEEWHIOYDEYNAPYGNNDKIWGYDDLASISCKLAFKELGVSQVWWSLE 360
DB 320 ELQCFQKEEWHIOYDEYNAPYGNNDKIWGYDDLASISCKLAFKELGVSQVWWSLE 379
QY 361 NDDKRGHCGPKNPLKVNMMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPT 420
DB 380 NDDKRGHCGPKNPLKVNMMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPT 439
QY 421 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
DB 440 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
QY 481 VDGLIKYKCEGDI PHTPTNHIKYLCEYFVNGWVWHIMPCPGTIWCEKLTGICE 536
DB 500 VDGLIKYKCEGDI PHTPTNHIKYLCEYFVNGWVWHIMPCPGTIWCEKLTGICE 555

RESULT 2
O44079
ID O44079 PRELIMINARY; PRT; 525 AA.
AC O44079;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Chitinase.
GN AGCHI-1.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gut;
RX MEDLINE=98030563; PubMed=9360958;
RA Shen Z., Jacobs-Lorena M.;
RT "Characterization of a novel gut-specific chitinase gene from the
RT human malaria vector Anopheles gambiae";
RL J. Biol. Chem. 272:28895-28900(1997).
DR EMBL; AF008575; AB87764.1; --
DR FIR; T4445; T4445.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006030; P:chitin metabolism; IEA.
DR InterPro; IPR002557; Chitin bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CSM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 1.

DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 525 AA; 57211 MW; 3234360EPPF36165 CRC64;
Query Match 33.4%; Score 1007; DB 5; Length 525;
Best Local Similarity 39.5%; Pred. No. 5.2e-55;
Matches 210; Conservative 86; Mismatches 178; Indels 58; Gaps 12;
QY 6 HNDYKNPVRIVCVGTWTSVYHKVDPTTIEDIDPFKCTHLMYGFADIKDYKTYIQVDP 64
DB 23 HKAASAEKGVYCVGTWAVYRPGNGRYDIEHIDPSLCTHLMYGFAGINE-DATVRIIDP 81
QY 65 YQDNNHNSWEKRGYERFNNLRLKNPELTMTISLGGWYEGSEKYSMDAANTPTYRQOF 124
DB 82 YLDLEN-WGRGHIKRFVGLKKNVGPGLKTLAIGWNEGRKESAAASGELRKRISDC 140
QY 125 LDFLQBYKFDGLDLDWEYFSGSRGNPKIDKQNYLALVRELKDAPEPHGYLLTAASVPGKD 184
DB 141 VAFQCRHGFDGLDLDWEYFSGSRGNPKIDKQNYLALVRELKDAPEPHGYLLTAASVVEF 200
QY 185 KIDRAYDIKELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNNTYWH 244
DB 201 SAGVSVDIIPRISKSPFLNVMYVDMEGAWDSYGINAPLYRGSADTTDLGQINNVASIH 260
QY 245 YLLNNGATRDKLVMGVPPFYGRAWSIEDRSKLKLGDPKAGMSPPGFISGEGVLSYIELCQ 304
DB 261 FWLAQCTGRKLVIGPLIGRNFLLASAANTQIGAPTGGTGVGRVTRPFGWGYNEFCE 320
QY 305 LFOKEBWHIOYDEYNAPYGNNDKIWGYDDLASISCKLAFKELGVSQVWWSLENDPF 364
DB 321 KLATEAWDLRWSBEQQVPYAVRNNQWGYDDLSVQLKYLDDQGLGAMWWSLETDDF 380
QY 365 KGHG-GPKNPLKVNMMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPTTPT 423
DB 381 LGVCGGGRYPLMHEIRSLVNGGT-----PSITTMPPSVAPT-----STVAPGTTT 426
QY 424 TTPSTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 483
DB 427 TPTGANPGTQPPT--SDAPNHTTSTTTEGNGPGRTPSPG-----DG 468
QY 484 HLIKCY--KEGDIPHTNHIKYLVC-----EFVNGWVWHIMPCPGTIW 526
DB 469 ---PCAGRGYGVFPHPTNCAARYICLTADTYEFT-----CPGTLF 507

RESULT 3
Q8ISH5
ID Q8ISH5 PRELIMINARY; PRT; 431 AA.
AC Q8ISH5;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Chitinase.
OS Araneus ventricosus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Araneioidea; Araneidae; Araneus.
OX NCBI_TaxID=182803;
RN [1]
RP SEQUENCE FROM N.A.
RA Han J.H., Kim S.R., Sohn H.D., Jin B.R.;
RT "Molecular cloning of a cDNA encoding the chitinase from the spider,
RT Araneus ventricosus";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY120879; AN39100.1; --
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.

[illegible]

```

227 PDETEHLHYFNWNYTHYLLNNGATKDLVMGVPFYIGRAWSJEDRSKLGKLDGPAKGMSP 286
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
880 PDDPDKT---FNANFTIHYWIEKGADPRKLVGMGMPYQGSFLADNKEHGLNAKTYGGGE 936
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
287 PGIISGEGVLSIELCOLFQKEWHIODEYN-APYGVNDKIWIYGVDDLASISCKLAF 345
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
937 AGESTRARGFLSYEICANIRNKKWTIVARDKGRMGFPYAKGQWVSFDDQVVMIRHKSEY 996
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 LKELGVGVGVMSLENDDFKHCYG-PKNPLLNKVHNMINGDEKNSPECIL--GPTSTTPT 402
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
997 VKAMGLGGAMIWALDDDFNLDCOEYPLLRINVLNRNYPGPGRCVLEKEPQREPR 1056
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
403 PTTTPTPTTTPTTPTSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 462
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1057 PTAPPIS-----TTPETTERSTTSTRITWTITTT--TTTRRPTTTRRTTSAR 1104
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 PSPTTTEHTSETPKY-----TTYVDGHLIKCYKEGDIPHPTNIHKYLVCSEFVNGW 513
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1105 PTYTIT---NVFSYQEIANEVDPECTDGRLF-----VPHTDCNKYIICQYGK--- 1150
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
514 WWHIMPCPPGTIW-----CQEKLT 532
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1151 -----LCPGGLYMSVDHCDWPQSTNCRNKQT 1176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
Q9W223 PRELIMINARY; PRT; 4498 AA.
AC Q9W223;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG2989 protein.
GN CG2989.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI TaxID=7227.

```



```

DR SMART; SM00494; ChtBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
SQ SEQUENCE 467 AA; 51958 MW; 734830C6F47F4CD CRC64;

Query Match
Best Local Similarity 27.6%; Score 831.5; DB 5; Length 467;
Matches 180; Conservative 87; Mismatches 174; Indels 85; Gaps 14;

QY 16 IVCVGTWSTYHK-VDPYTTIEDIPFKCTHLMYGFPAKIDYKTIQVDFYQD--DNHNS 72
Db 1 MCVYFGSWAAYRQGLGKDFVEDIPKCTHIIIFGAGL-AHDSIRVLDPNWLCDNYG- 58
QY 73 WEKGYERFNRLKKNPELATTMSLGGWYEGSEKYSMDAANPYRQOFTQSLVDFLOEYK 132
Db 59 --KCAVDRFTALKOONANLKAILAVGWNNEGSPKYSKMAADPVLRNRTFSS-TELLKGG 116
QY 133 FDGLDLWEYFGSLGNPKDKONYLALVRELKDAPEPHGYLLTAAVSPGKIDRAYDI 192
Db 117 FDGLDMWEYFTQGGSPD-DYDNFVILMAELNQAALHAEGMLTAAVSAGATIDPAYV 175
QY 193 KELNKLFDWNNVMTYDHGGWENFYGHNAPLYKRP-DETDLTFTFNVTMYHYLNNGA 251
Db 176 PELSKSLDLNVMYTDLHGAWDDVTHQSGLYAHPLDEGD--NAFLNVDPFAISYWIEKA 233
QY 252 TRKLNVGVFPYGRANSIEDRSKLGDPAKWSPPGFISGERGVLSYELCOLFOKEW 311
Db 234 RPGQIALGILYGRCWTLASQETGYAPAHQPCAAGDWTSPGLGYNEICVMYRTQDW 293
QY 312 HIQVDEYVNAFYGI---NDKIWGYDLDASISCKLAFKELGVLGVNWSLENDPFGHC 368
Db 294 TVVDDPANNFEYTFYFPMNNINGSDYHAASVAIKAEYAKSGLAGTMMWSVEFDDFRGC 353
QY 369 GPKNPLINKVHNMINGDEKNSFCILGPTTTTPTTPTTPTTPTTPTTPTTPTTPTTSP 428
Db 354 H-----NRKYHLI-----KTMVEFGGGSITEP----- 376
QY 429 TPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 488
Db 377 -----PLPPTTRDNEPTTTRAPP-----PG-----VHC 403
QY 489 YKGGDIPHPNINIKLVCEF-VNGGWVHIMPCTGTCINQCKLTC 533
Db 404 TQPLNPDPLDCTHYLCSLNTSGYNEKEEVCPEGLVNPQSYIC 449

RESULT 10
Q9W2M7 PRELIMINARY; PRT; 460 AA.
AC Q9W2M7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG9357 protein.
GN CG9357.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.C., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```


DR GO: 0008843; F: endochitinase activity; IEA.
 DR GO: 0016798; F: hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO: 0005975; P: carbohydrate metabolism; IEA.
 DR GO: 0006030; P: chitin metabolism; IEA.
 DR InterPro: IPR002557; Chitin bind PerA.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR InterPro: IPR001579; Glyco_hydro_18AS.
 DR Pfam: PF01607; CEM_14; 1.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR ProDom: PD000471; Glyco_hydro_18; 1.
 DR SMART: SM00494; ChitBD2; 1.
 DR SMART: SM00636; Glyco_18; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 DR GLYCOBASE: Hydrolase; Signal.
 KW SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 543 CHITINASE.
 FT SEQUENCE 543 AA; 60971 MW; C36C71E3B84DF88 CRC64;
 SQ
 Query Match 27.2%; Score 819; DB 5; Length 543;
 Best Local Similarity 36.6%; Pred. No. 3e-43;
 Matches 200; Conservative 72; Mismatches 203; Indels 72; Gaps 18;
 QY 10 SKNPMRIVCYVGTWSVYHK-VDPYTIEDIDPKCTHLMYGFADKIDYKTIQVDFPVQDD 68
 DB 20 SDSRARIVCYFSNVAWYRPGVGRYGIEDIPVLCITLIYSFIVGTEKSSSEVLIDPELD- 78
 QY 69 NHNSWEKRGYRFFNNLRKLNKPELTMTISLGWYEGSEKYSMDAANPYRQOFTQSVLDLFL 128
 DB 79 ----VDKSGPRFTSLRSKHPDVKFMVAVGWAEGGSKYSHWAQKSTRMSFIRSVDFL 134
 QY 129 QEYKFDGLDLDWEYPCGR-LGNPKIDKQNYLALVRELKDAF--EPHGYLITAAVSPGKDK 185
 DB 135 KKYDFDGLDLDWEYPCGR-LGNPKIDKQNYLALVRELKDAF--EPHGYLITAAVSPGKDK 194
 QY 186 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKRPDETDELHLYFNVTMYH 245
 DB 195 LMGGYHVPCLQELDAIHVMSYDLRGWAGFADVHSPLYKRPD-QWAYEKLNVNDGLNL 253
 QY 246 YLNNGATRDKLVMGVPFYGRANSIED-----RSKLKLDGPAKMSPPGFTSGEE 294
 DB 254 WEEKGCPNKLIVGIPFGYGRSFTLSAGNNYGLGTVINKEAGGDPAPYTNATGF----- 308
 QY 295 GVLSYIELCOLFOKE--EWHIOYDEVYNAPYGVNDKIWGVDDLASISCKLAFKLKELGVS 352
 DB 309 --WAYEICTEVDADGSGWKWDEFGKCPYAKGTQWVGVEDPRSVVEIKMWIKKGYL 366
 QY 353 GVMVWSLENDDFKGGCKPNLKNVHNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTT 412
 DB 367 GANTWALDMDDFKGLCGSENLKLLRHMS-----THTVPPARTGHTTPTT 413
 QY 413 TPTTPTTPTTPTT-----TPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 468
 DB 414 EWARP-PSTPSPDSEGDPIPTTTTIVKPTTTRTARPTT-----TTTKVPHGTTT 463
 QY 469 E-----HTSETPKYTYVDGHLIKCYKEGD-IPHPTNIHKYLVCFVNGGWWHIMP 519
 DB 464 EDFDINVRPEVEELPT-ENEVDNADV-CNSEDDYVDPKCKESKYWRC--VNGEGVQ--FS 517
 QY 520 CPPTGIM 526
 DB 518 CQPGTIF 524
 RESULT 15
 Q9GV05 PRELIMINARY; PRT; 543 AA.
 AC Q9GV05;
 DT 01-WAR-2001 (TrEMBLrel. 16, Created)
 DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Chitinase precursor (BC 3.2.1.14).
 GN CHI.
 OS Bombyx mori (Silk moth).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kinsu X Showa hybrid; TISSUE=Integument;
 RX MEDLINE=21124800; PubMed=1122960;
 RA Abdel-Banat B.M.A.; Koga D.;
 RT "A genomic clone for a chitinase gene from the silkworm, Bombyx mori:
 RL structural organization identifies functional motifs.";
 RL Insect Biochem. Mol. Biol. 31:497-508(2001).
 DR EMBL; AB048355; BAB13481.1; .
 DR GO: 0005576; C: extracellular; IEA.
 DR GO: 0008061; F: chitin binding; IEA.
 DR GO: 0008843; F: endochitinase activity; IEA.
 DR GO: 0016798; F: hydrolase activity, acting on glycosyl bonds; IEA.
 DR GO: 0005975; P: carbohydrate metabolism; IEA.
 DR GO: 0006030; P: chitin metabolism; IEA.
 DR InterPro: IPR002557; Chitin bind PerA.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR InterPro: IPR001579; Glyco_hydro_18AS.
 DR Pfam: PF01607; CEM_14; 1.
 DR ProDom: PD000471; Glyco_hydro_18; 1.
 DR SMART: SM00494; ChitBD2; 1.
 DR SMART: SM00636; Glyco_18; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 543 CHITINASE.
 FT SEQUENCE 543 AA; 60982 MW; DE29675D83AEC2EF CRC64;
 SQ
 Query Match 27.2%; Score 819; DB 5; Length 543;
 Best Local Similarity 36.6%; Pred. No. 3e-43;
 Matches 200; Conservative 72; Mismatches 203; Indels 72; Gaps 18;
 QY 10 SKNPMRIVCYVGTWSVYHK-VDPYTIEDIDPKCTHLMYGFADKIDYKTIQVDFPVQDD 68
 DB 20 SDSRARIVCYFSNVAWYRPGVGRYGIEDIPVLCITLIYSFIVGTEKSSSEVLIDPELD- 78
 QY 69 NHNSWEKRGYRFFNNLRKLNKPELTMTISLGWYEGSEKYSMDAANPYRQOFTQSVLDLFL 128
 DB 79 ----VDKSGPRFTSLRSKHPDVKFMVAVGWAEGGSKYSHWAQKSTRMSFIRSVDFL 134
 QY 129 QEYKFDGLDLDWEYPCGR-LGNPKIDKQNYLALVRELKDAF--EPHGYLITAAVSPGKDK 185
 DB 135 KKYDFDGLDLDWEYPCGR-LGNPKIDKQNYLALVRELKDAF--EPHGYLITAAVSPGKDK 194
 QY 186 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPLYKRPDETDELHLYFNVTMYH 245
 DB 195 LMGGYHVPCLQELDAIHVMSYDLRGWAGFADVHSPLYKRPD-QWAYEKLNVNDGLNL 253
 QY 246 YLNNGATRDKLVMGVPFYGRANSIED-----RSKLKLDGPAKMSPPGFTSGEE 294
 DB 254 WEEKGCPNKLIVGIPFGYGRSFTLSAGNNYGLGTVINKEAGGDPAPYTNATGF----- 308
 QY 295 GVLSYIELCOLFOKE--EWHIOYDEVYNAPYGVNDKIWGVDDLASISCKLAFKLKELGVS 352
 DB 309 --WAYEICTEVDADGSGWKWDEFGKCPYAKGTQWVGVEDPRSVVEIKMWIKKGYL 366
 QY 353 GVMVWSLENDDFKGGCKPNLKNVHNMINGDEKNSFECILGSPSTTTPTTPTTPTTPTT 412
 DB 367 GANTWALDMDDFKGLCGSENLKLLRHMS-----THTVPPARTGHTTPTT 413
 QY 413 TPTTPTTPTTPTT-----TPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 468
 DB 414 EWARP-PSTPSPDSEGDPIPTTTTIVKPTTTRTARPTT-----TTTKVPHGTTT 463
 QY 469 E-----HTSETPKYTYVDGHLIKCYKEGD-IPHPTNIHKYLVCFVNGGWWHIMP 519
 DB 464 EDFDINVRPEVEELPT-ENEVDNADV-CNSEDDYVDPKCKESKYWRC--VNGEGVQ--FS 517

QY 520 CQPGTIW 526
| | | | |
Db 518 CQPGTIF 524

Search completed: March 22, 2004, 06:59:27
Job time : 108.855 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:31:13 ; Search time 20.5581 Seconds
(without alignments)
1357.597 Million cell updates/sec

Title: US-09-662-293-21
Perfect score: 3014
Sequence: 1 SIKEDHNDYKKNPRIVCYV.....IMCPPTGTWCQKLTICGE 536

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 834 | 27.7 | 554 | 1 | CHIT_MANSE |
| 2 | 738 | 24.5 | 466 | 1 | CHIT_HUMAN |
| 3 | 727.5 | 24.1 | 473 | 1 | CHIA_MOUSE |
| 4 | 723 | 24.0 | 508 | 1 | CHIL_DROME |
| 5 | 708.5 | 23.5 | 476 | 1 | CHIA_HUMAN |
| 6 | 705.5 | 23.4 | 617 | 1 | CHIT_CAEEL |
| 7 | 694.5 | 23.0 | 527 | 1 | OGP_PIG |
| 8 | 682.5 | 22.6 | 531 | 1 | OGP_SHEEP |
| 9 | 680 | 22.6 | 381 | 1 | C3LI_MOUSE |
| 10 | 680 | 22.6 | 721 | 1 | OGP_MOUSE |
| 11 | 677.5 | 22.5 | 671 | 1 | OGP_MESAU |
| 12 | 673 | 22.3 | 537 | 1 | OGP_BOVIN |
| 13 | 667.5 | 22.1 | 678 | 1 | OGP_HUMAN |
| 14 | 660 | 21.9 | 383 | 1 | C3LI_HUMAN |
| 15 | 654 | 21.7 | 623 | 1 | OGP_PAPAN |
| 16 | 630 | 20.9 | 390 | 1 | C3L2_HUMAN |
| 17 | 611.5 | 20.3 | 396 | 1 | CHIT_MOUSE |
| 18 | 604.5 | 20.1 | 504 | 1 | CHIT_BRUMA |
| 19 | 576 | 19.1 | 458 | 1 | CHIT_DROME |
| 20 | 561 | 18.6 | 699 | 1 | CHIL_BACCI |
| 21 | 409.5 | 13.6 | 423 | 1 | CHIL_APHAL |
| 22 | 407.5 | 13.5 | 563 | 1 | CHIA_SERVA |
| 23 | 402.5 | 13.4 | 423 | 1 | CHIA_TRIHA |
| 24 | 384 | 12.7 | 820 | 1 | CHIA_ALTISO |
| 25 | 377.5 | 12.5 | 550 | 1 | CHIT_NPVOP |
| 26 | 360.5 | 12.0 | 551 | 1 | CHIT_NPVAC |
| 27 | 353 | 11.7 | 499 | 1 | CHIB_SERVA |
| 28 | 341 | 11.3 | 619 | 1 | CHIT_STRLI |
| 29 | 339.5 | 11.3 | 427 | 1 | CHIL_COEPO |
| 30 | 326 | 10.8 | 610 | 1 | CHIT_STRPL |
| 31 | 280 | 9.3 | 1046 | 1 | CHID_VIBFU |
| 32 | 266.5 | 8.8 | 474 | 1 | VTP3_TTVIV |
| 33 | 259 | 8.6 | 5179 | 1 | MUC2_HUMAN |

| | | | | | |
|----|-------|-----|------|---|------------|
| 34 | 247 | 8.2 | 1513 | 1 | MUC1_RAT |
| 35 | 244 | 8.1 | 662 | 1 | MUC1_XENLA |
| 36 | 238 | 7.9 | 279 | 1 | Y091_NFVOP |
| 37 | 230 | 7.6 | 400 | 1 | MUAI_XENLA |
| 38 | 227.5 | 7.5 | 1331 | 1 | MANB_CALSA |
| 39 | 221 | 7.3 | 267 | 1 | EXTN_MAIZE |
| 40 | 217.5 | 7.2 | 544 | 1 | GP10_DICDI |
| 41 | 217.5 | 7.2 | 1161 | 1 | DAN4_YEAST |
| 42 | 215.5 | 7.1 | 705 | 1 | GUN6_DICDI |
| 43 | 215.5 | 7.1 | 1146 | 1 | KTXA_KLULA |
| 44 | 208.5 | 6.9 | 283 | 1 | EXTN_SORBI |
| 45 | 207 | 6.9 | 307 | 1 | SGS3_DROME |

ALIGNMENTS

RESULT 1

| ID | CHIT_MANSE | STANDARD; | PRT; | 554 AA. |
|----|---|-----------------------------------|------|---------|
| DT | 01-JUN-1994 | (Rel. 29, Created) | | |
| DT | 01-JUN-1994 | (Rel. 29, Last sequence update) | | |
| DT | 15-MAR-2004 | (Rel. 43, Last annotation update) | | |
| DE | Endochitinase precursor (EC 3.2.1.14). | | | |
| OS | Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; | | | |
| OC | Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea; | | | |
| OC | Sphingidae; Sphinginae; Manduca. | | | |
| OX | NCBI_TaxID=7130; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=93357793; PubMed=8353525; | | | |
| RA | Kramer K.J., Corpuz L., Choi H.K., Muthukrishnan S.; | | | |
| RT | "Sequence of a cDNA and expression of the gene encoding epidermal and gut chitinases of Manduca sexta." | | | |
| RL | Insect Biochem. Mol. Biol. 23:691-701(1993). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=97215580; PubMed=9061927; | | | |
| RA | Choi H.K., Choi K.H., Kramer K.J., Muthukrishnan S.; | | | |
| RT | "Isolation and characterization of a genomic clone for the gene of an insect molting enzyme, chitinase." | | | |
| RL | Insect Biochem. Mol. Biol. 27:37-47(1997). | | | |
| CC | !- FUNCTION: Digest chitin in the exoskeleton during the molting process. | | | |
| CC | !- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-acetyl-D-glucosamine polymers of chitin. | | | |
| CC | !- SUBCELLULAR LOCATION: Secreted. | | | |
| CC | !- TISSUE SPECIFICITY: Epidermis and gut. | | | |
| CC | !- DEVELOPMENTAL STAGE: High levels seen in the epidermis on day 0, but rapidly disappears and is undetected on days 1-4 of fifth instar. It reappears on day 5 and peaks on day 7 after which a rapid decline is seen. In the gut is detected on day 6 with lower levels seen on days 0, 7 and 8. | | | |
| CC | !- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl hydrolases). | | | |
| CC | !- SIMILARITY: Contains 1 chitin-binding type-2 domain. | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |
| DR | EMBL; U02270; AAC04924.1; - | | | |
| DR | EMBL; L49234; AAB53952.1; - | | | |
| DR | PIR; A56596; A56596 | | | |
| DR | InterPro; IPR002557; Chitin_bind_Pera. | | | |
| DR | InterPro; IPR001223; Glyco_Hydro_18. | | | |
| DR | InterPro; IPR001579; Glyco_Hydro_18AS. | | | |

```
DR Pfam; PF01607; CBM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChitBD2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS00940; CHIT_BIND_II; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 554 ENDOCHITINASE.
FT DOMAIN 396 453 SER/THR-RICH.
FT DOMAIN 495 553 CHITIN-BINDING TYPE-2.
FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 554 AA; 62203 MW; 3989D756C96CD490 CRC64;

Query Match 27.7%; Score 834; DB 1; Length 554;
Best Local Similarity 35.9%; Pred. No. 1.3e-38;
Matches 203; Conservative 76; Mismatches 202; Indels 84; Gaps 19;

QY 10 SKNMRIVCYVGTWGVYHK-VDPYTIEDIDPFKCTHLMYGFADIDEXYTIQVDFPYQDD 68
Db 19 SDSRARIYCFENWAVYRGVRYGIEDIPVEKCTHIYISFVIGTEGSEVLIDPELD- 77
QY 69 NNSWEKGYEFNNLRKNPELTMTISLGGWEGSEKYSMAANTYRQOFIOSVDPL 128
Db 78 ---VDKNGFRNFTSLRSHSPKFWAVGWAEGSKYSHWVAKQSRMSFIRSVSFL 133
QY 129 QBYKEDGLDLDWEYQSGR-LGNPKYDKQNYLALVPELKDAP--EPHGVLTLTAAYSPGKX 185
Db 134 KYDFDGLDLDWEYFCAADRGSGFSFDKFLYVQELRRATIRVGKQWELTAVPLANFR 193
QY 186 IDRAVDIKELNKLFDWNVMTYDYGWENFYGHNAPLYKRPDETDLHTFNVNVTYHY 245
Db 194 LMEGHVPELCELDIAHVMSTDLRGNWAGFADVHSPLYKRPD-QWYAEKLVNDGHL 252
QY 246 YLNGATRDKLVGVFPFYGRAWSIED-----RSKLGLGPAKMGSPPGFISGEE 294
Db 253 WEKGCPSNKLNVGIPFYGRSFTLSAGNNVGLGTFINKEAGGDPAYTNATGF----- 307
QY 295 GVLSYIELCOLFOKEE--WHIYDEYNAPYNDKIWGYDDIASISCKLAFKLGLVS 352
Db 308 --WAYVEICTEVDKDSGWTKWDEGCKPYAYKGTQWVGVEDPDRSVBIKNNWIKQGYL 365
QY 353 GYVWVSLNDPKGHGCPKNPLLNKVHNMINGDERKNSFECILGPSTTTTPTTPTTPT- 411
Db 366 GANTWAIMDDFQGLGCKGNPLIKLTHKMS-----SYTVPPPHHTNTTPTTPT 412
QY 412 -----TTPTPS-----PTTPTPSPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 450
Db 413 ENWRPSTSDSEGDPIPTTTTAKPASTKTIVKTTTAKPQSVIDEENDINVRPE 472
QY 451 P-TTPTAPTSTPSPTTTEHTSETPKYTYVDGHLIKCYKGD-IPHPTNIHKVLVCEP 508
Db 473 PKPEQPEPEVEVP-PTNE-----VDGSEI-CNSDDYIPDKKHCDKYRC-- 517
QY 509 VNGGWVHMLNCPGPTIWCQEKUTC 533
Db 518 VNGE--AMQFSCQGTGVFNELNVC 540

RESULT 2
CHIT_HUMAN
ID CHIT1_HUMAN STANDARD; PRT; 466 AA.
AC Q13231; Q9H3V8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chitotriosidase 1 precursor (EC 3.2.1.14) (Chitinase 1).
```

```
GN CHIT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND SUBCELLULAR
RP LOCATION.
RC TISSUE=Macrophage;
RX MEDLINE=96064695; PubMed=7592832;
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J.,
RA Aerts J.M.F.G.;
RT "Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT produced by macrophages."
RL J. Biol. Chem. 270:26252-26256(1995).
RN [2]
RP SEQUENCE OF 22-42 AND 178-198, FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=95138187; PubMed=7836450;
RA Renkema G.H., Boot R.G., Muijsers A.O., Donker-Koopman W.E.,
RA Aerts J.M.F.G.;
RT "Purification and characterization of human chitotriosidase, a novel
RT member of the chitinase family of proteins."
RL J. Biol. Chem. 270:2198-2202(1995).
RN [3]
RP POLYMORPHISM, AND DESCRIPTION OF ISOFORM 3.
RX MEDLINE=98421482; PubMed=9748235;
RA Boot R.G., Renkema G.H., Verhoek M., Strijland A., Bliet J.,
RA de Meuleneester T.M.A.M.O., Mannens M.M.A.M., Aerts J.M.F.G.;
RT "The human chitotriosidase gene. Nature of inherited enzyme
RT deficiency."
RL J. Biol. Chem. 273:25680-25685(1998).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 22-386 IN COMPLEX WITH
RP CHITOBIOSE AND ALLOSAMIDIN.
RX MEDLINE=22095530; PubMed=11960986;
RA Fusetti F., von Moeller H., Houston D., Rozeboom H.J., Dijkstra B.W.,
RA Boot R.G., Aerts J.M.F.G., van Aalten D.M.F.;
RT "Structure of human chitotriosidase. Implications for specific
RT inhibitor design and function of mammalian chitinase-like lectins."
RL J. Biol. Chem. 277:25537-25544(2002).
CC -!- FUNCTION: Degrades chitin and chitotriose. May participate in the
CC defense against nematodes and other pathogens. Isoform 3 has no
CC enzymatic activity.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted. A small proportion is lysosomal.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q13231-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q13231-2; Sequence=VSP_008631, VSP_008632;
CC Name=3;
CC IsoId=Q13231-3; Sequence=VSP_008633;
CC Note=Duplication of 24 bp in exon 10 leads to the use of a
CC cryptic splice site. The normal splice site is still present but
CC not used;
CC -!- TISSUE SPECIFICITY: Detected in spleen. Secreted by cultured
CC macrophages.
CC -!- POLYMORPHISM: A 24 bp duplication in exon 10 leads to the
CC activation of an alternative splice site and the production of an
CC inactive protein. About 6% of the population are deficient for
CC CHIT1 activity, while 35% are carriers and show reduced enzyme
CC levels. People with CHIT1 deficiency appear perfectly healthy.
CC -!- MISCELLANEOUS: Patients with type 1 Gaucher disease (GD I)
CC [MIM:230800] have very high plasma levels of CHIT1, and this can
CC be used as diagnostic aid and to evaluate the success of
CC treatment. Successful therapy brings the CHIT1 activity levels
CC back to normal.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
```

| | | | | | | | | | | | | |
|----|---|---|--|--|--|--|--|--|--|--|--|----|
| CC | CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | | | | | | | | CC |
| CC | EMBL; U25615; AAC50246.1; -- | CC | | | | | | | | | | |
| DR | EMBL; U62662; AGL0644.1; -- | DR | | | | | | | | | | |
| DR | PDB; 1GUU; 31-JAN-02. | DR | | | | | | | | | | |
| DR | PDB; 1LGI; 18-SEP-02. | DR | | | | | | | | | | |
| DR | PDB; 1LG2; 18-SEP-02. | DR | | | | | | | | | | |
| DR | PDB; 1LQ0; 29-JUL-03. | DR | | | | | | | | | | |
| DR | Genew; HGNC:1936; CHIT1. | DR | | | | | | | | | | |
| DR | MM; 600031; -- | DR | | | | | | | | | | |
| DR | GO; GO:0005615; C:extracellular space; TAS. | DR | | | | | | | | | | |
| DR | GO; GO:0004568; F:chitinase activity; TAS. | DR | | | | | | | | | | |
| DR | GO; GO:0009617; P:response to bacteria; TAS. | DR | | | | | | | | | | |
| DR | GO; GO:0009613; P:response to pest/pathogen/parasite; TAS. | DR | | | | | | | | | | |
| DR | InterPro; IPR002557; Chitin_bind_PeA. | DR | | | | | | | | | | |
| DR | InterPro; IPR001223; Glyco_Hydro_18. | DR | | | | | | | | | | |
| DR | InterPro; IPR001579; Glyco_hydro_18AS. | DR | | | | | | | | | | |
| DR | Pfam; PF01607; CBM_14; 1. | DR | | | | | | | | | | |
| DR | Pfam; PF00704; Glyco_hydro_18; 1. | DR | | | | | | | | | | |
| DR | ProDom; PD000471; Glyco_hydro_18; 1. | DR | | | | | | | | | | |
| DR | SMART; SM00494; ChtBD2; 1. | DR | | | | | | | | | | |
| DR | SMART; SM00636; Glyco_18; 1. | DR | | | | | | | | | | |
| DR | PROSITE; PS00940; CHIT_BIND_II; 1. | DR | | | | | | | | | | |
| DR | PROSITE; PS01095; CHITINASE_18; 1. | DR | | | | | | | | | | |
| DR | Carbohydrate metabolism; Chitin degradation; | DR | | | | | | | | | | |
| KW | Polyaccharide degradation; Hydrolase; Glycosidase; Chitin-binding; | KW | | | | | | | | | | |
| KW | Signal; Alternative splicing; 3D-structure. | KW | | | | | | | | | | |
| FT | SIGNAL 1 21 | FT | | | | | | | | | | |
| FT | CHAIN 22 466 | FT | | | | | | | | | | |
| FT | DOMAIN 417 466 | FT | | | | | | | | | | |
| FT | ACT_SITE 140 140 | FT | | | | | | | | | | |
| FT | DISULFID 26 51 | FT | | | | | | | | | | |
| FT | DISULFID 307 370 | FT | | | | | | | | | | |
| FT | VARSPLIC 386 387 | FT | | | | | | | | | | |
| FT | VARSPLIC 388 466 | FT | | | | | | | | | | |
| FT | VARSPLIC 344 372 | FT | | | | | | | | | | |
| FT | SEQUENCE 466 AA; 51681 MW; B4312D1E883E386D CRC64; | FT | | | | | | | | | | |
| QY | Query Match 24.5%; Score 738; DB 1; Length 466; | QY | | | | | | | | | | |
| Db | Best Local Similarity 34.28; Pred No. 1.8e-33; | Db | | | | | | | | | | |
| QY | Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11; | QY | | | | | | | | | | |
| Db | 15 RIVCYGTGWSVYKVDP-YTIEDIDPKCTHLMYGFAKIDYEYKTIQVDFDPQDDNHSW 73 | Db | | | | | | | | | | |
| QY | 23 KLCYFTNWAQYRQGEARFLPKOLDPSLCTHLIVAFAGMTHQLST-----TEW 71 | QY | | | | | | | | | | |
| Db | 74 -EKRGVRFNNRLKNPELTMISLQGWYEGSEKYSDMAANPTVROQFIQSVLDFLOEYK 132 | Db | | | | | | | | | | |
| Db | 72 NDELTVEGFNGLKQNPKLKTLTAIGCGWNTGTQKFTDWATTANNRQFFVNSAIRFLRKY 131 | Db | | | | | | | | | | |
| QY | 133 FDGLDLDMWYPSGSELGNPKDKQNYLALVRELKDAPEFHG-----YLITAAVSPGDK 185 | QY | | | | | | | | | | |
| Db | 132 FDGLDLDMWYPSGQ-GSPAYDKERFTTLYQDLANAFQOEAQTSGKERLLLSAAVPAGQY 190 | Db | | | | | | | | | | |
| QY | 186 IDRAYDIKEINLKLFDWNNVMTYDHYGQWENFYGHNAPLYKRPDETDELHYFVNNYTMHY 245 | QY | | | | | | | | | | |
| Db | 191 VDAQYEVDKIAQNLDLFVNLWAYDFHGWSEKVTGHSPLYKRQESGAAS-LNVDAAVQQ 249 | Db | | | | | | | | | | |
| QY | 246 YLANKGATRDKLVMGVPPFYGAWSTEDSKLKLGDPAKGMSPGPGFISGEEGVLSYIELCQL 305 | QY | | | | | | | | | | |
| Db | 250 WLQKGFTASKLLIGMPTFYGHSFTLIASSSDTRVGAPATGSGTGPFTKEGMLAYEVCVS 309 | Db | | | | | | | | | | |
| QY | 306 FQKEWHIQDYDEYNAPYGNDKITWGYDDLASISCKLAFIKELGLGVSQVMVWLENDDFK 365 | QY | | | | | | | | | | |

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettenan M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 2-473 FROM N.A.
RC TISSUE-Skin;
RA Price P.A., Harris S.C., Williamson M.K.;
RT "YnL, a putative mouse chitinase.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Degrades chitin and chitotriose. May participate in the
CC defense against nematodes and other pathogens.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in submaxillary gland, and
CC stomach. Expressed at lower levels in lung.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF290003; BAG60018.1; -;
DR EMBL; AK008633; BAB25795.1; -;
DR EMBL; BC011134; RAH11134.1; ALT INIT.
DR EMBL; BC034548; RAH34548.1; ALT INIT.
DR EMBL; AF154571; AAF31644.1; -;
DR MGD; MGI:1932052; Chia.
DR GO; GO:0004568; F:chitinase activity; IDA.
DR GO; GO:0006032; P:chitin catabolism; IDA.
DR InterPro; IPR002557; Chitin_bind_Pera.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF01607; CSM_14; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtB2; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS50940; CHIT_BIND_II; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR Carbohydrate metabolism; Chitin degradation;
KW Polysaccharide degradation; Hydrolase; Glycosidase; Chitin-binding;
KW Signal.
FT SIGNAL 1 21
FT CHAIN 22 473 ACIDIC MAMMALIAN CHITINASE.
FT DOMAIN 424 473 CHITIN-BINDING TYPE-2.
FT ACT_SITE 140 140 BY SIMILARITY.
FT DISULFID 26 51 BY SIMILARITY.
FT DISULFID 307 372 BY SIMILARITY.
FT CONFLICT 293 P -> A (IN REF. 1).
SQ SEQUENCE 473 AA; 52003 MW; 333C874477476695 CRC64;
Query Match 24.1%; Score 727.5; DB 1; Length 473;
Best Local Similarity 37.0%; Pred. No. 6.7e-33;
Matches 154; Conservative 69; Mismatches 144; Indels 49; Gaps 11;
QY 16 IVCYVGTWSVYHK-VDPTIEDIPFKCTHLMYGFADKIDEXYKTIQVDFPDQDDNHNWS- 73
DB 24 LICYFTNWAQYRPLGSGFKPDDINPCLTCTHLYAFAGMKNNEIT- -----IENW 72

QY 74 EKRGYERFNNLRKLNPELTMTISLGWYEGSEKYSDMAANPTVROQFIQSILDFLOEYKF 133
DB 73 DVTLYKAFNDLKNRNSKLTLLAIGWNTGTAPFTTWSTQNRQTFITSVIKFLQYGF 132
QY 134 DGLDLWEYPSGSLGNPKIDKQNYLALVELKDAFEPHG-----YLTAAVSPGKDKI 186
DB 133 DGLDLWEYPSGR-GSPPODKHLFTVLKEMREAPQEAIESNRPLMTVAAGGISNI 191
QY 187 DRAYDKELNKLFDWVWNTYDYHGWNFYGHNAPLYKRPDETDLHTYFNVNTHYY 246
DB 192 QAGYEIPELSKYLDYFHVNTYDLHGSWEGYTGNSLYKPYTETGS-NAYLVNDYVWY 250
QY 247 LKNGATEDKLVMGVPPFYGRAWSIEDRSKLGDKPAKMSPPGPGFISGEEGVLSEICOLF 306
DB 251 KXNGAPAEKLVGVFPYGHFTILRNPSDNGIGAPTSGDGPAGYTRQAGFWAYEICTFL 310
QY 307 QK---BEWHIQYDEYVNPAGYNDKIKWGVYDDIASISCKLAFKELGLVSGVMMVSLNDD 363
DB 311 RSGATEVW----DASQEVFVAYKANELGVDNIKSFVKAQWLKQNNFGGAMTWAIDDD 366
QY 364 FKQ-HCG-----PKNPLLNKVNMMINGDEKNSFECILGPSTTTPTTTPTTTPTTT 414
DB 367 FTGSCFCDQKGFPLTSTLNKA-----LGISTEGCTAPDVPSEPTTP 407
RESULT 4
CH11 DROME STANDARD; PRT; 508 AA.
ID CH11 DROME STANDARD; PRT; 508 AA.
AC Q9W5J3; O17420;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable chitinase 1 (EC 3.2.1.14).
GN CHT1 OR CG17682.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=22426071; PubMed=12537574;
RA Hoskins R.A., Smith C.D., Carlson J.W., Carvalho A.B., Halpern A.,
RA Kaminker J.C., Kennedy C., Mungall C.J., Sullivan B.A., Sutton G.G.,
RA Yasuhara J.S., Wakimoto B.T., Myers E.W., Celniker S.E., Rubin G.M.,
RA Karpen G.H.;
RT "Heterochromatic sequences in a Drosophila whole-genome shotgun
RT assembly.";
RL Genome Biol. 3:RESEARCH0085.1-RESEARCH0085.16(2002).
RN [2]
RP SEQUENCE OF 151-263 FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=98324849; PubMed=9662472;
RA de la Vega H., Specht C.A., Liu Y., Robbins P.W.;
RT "Chitinases are a multi-gene family in Aedes, Anopheles and
RT Drosophila.";
RL Insect Mol. Biol. 7:233-239(1998).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl
CC hydrolases).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF026500; AAB81858.1; -;
DR EMBL; FBgn0022703; Chit1.

```
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 2.
DR ProDom; PD000471; Glyco_hydro_18; 2.
DR SMART; SM00836; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Multigene family.
FT ACT SITE 264 264 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 508 AA; 57751 MW; 26CA23802EFD97 CRC64;

Query Match
Best Local Similarity 39.8%; Score 723; DB 1; Length 508;
Matches 144; Conservative 71; Mismatches 121; Indels 26; Gaps 9;

QY 15 RIVYVGTWNVYK-VDPYTIEDIDPEKCTHLMYGAIDYKTYITQVDPQDDNNHNSW 73
DB 143 KIICYFTNWAYRQGGKFLPEDIISDLCTHIYGFVLSRDNLTIQ-----PHDSW 194
QY 74 --EKGYEFNNLRKNPELTTWISLGGWYEGS-EKYSDMAANPTVROFTQSVLDFLQ 129
DB 195 ADLNKPYEIVAYRKGAQVT--VAIGWNDSAGDKYSLVRNPEARSRFIRNVLDFIE 252
QY 130 EYKFDGLDLDEWYP-----GSRLEGNPKIDKQNTYALVRELKDAFEPHGYLLTAASPGKD 184
DB 253 EYNFDGLDLDEWYPCWQVDCCKGTAE-EKIGESALVRELFPYAFQPRGLILSNAVSPKK 311
QY 185 KIDRAYDIKELKLPDMNMYTYDYHGMENFYGHNAPLYKRPDETDLHTYFNVNVTWH 244
DB 312 VIDAGYEVAELSHVFNISWVAYDYGQWDKKTGHVAPVMSHEGT----ANFNANFSWN 367
QY 245 YLLNGATRLKLVGPPFYGRANSIEDRSKILGDPKAGMSPPGFTSGEGVLSYELCQ 304
DB 368 YWISMGADRKLKLVGPIYLGQSFSLAETTKHQLNAPTYGGGEAGEATRAGFLAYVEICL 427
QY 305 LFQKEEHHIQYDEVYN-APGYNDKTVGWYDDLASISKLAFLKEIVGSGVMVNSLEND 363
DB 428 KIRHRNVRVDTKGRIGPPAYHGQDQVSPDDVFNIRHKESEYIKAMGLGAMINWALDLD 487
QY 364 FK 365
DB 488 FK 489

RESULT 5
CHIA_HUMAN STANDARD; PRT; 476 AA.
AC Q9BZP6; Q86UD8; Q9ULY3; Q9ULY4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Acidic mammalian chitinase precursor (EC 3.2.1.14) (AMCase)
DE (TSA1902).
GN CHIA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Lung;
RX MEDLINE=20018184; PubMed=10548734;
RA Saito A., Ozaki K., Fujiwara T., Nakamura Y., Tanigami A.;
RT "Isolation and mapping of a human lung-specific gene, TSA1902, encoding a novel chitinase family member.";
RL Gene 239:325-331(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Lung, and Stomach.
RX MEDLINE=21125893; PubMed=11085997;
RA Root R.G., Blommaert E.F.C., Swart B., Ghauharali-van der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
RT "Identification of a novel acidic mammalian chitinase distinct from chitotriosidase.";
```

```
RL J. Biol. Chem. 276:6770-6778 (2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano P.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smal M.A.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Degrades chitin and chitotriose. May participate in the defense against nematodes and other pathogens.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-acetyl-D-glucosamine polymers of chitin.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable). Isoform 2 and isoform 3 lack the signal sequence and are cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9BZP6-1; Sequence=Displayed;
CC Name=2; Synonyms=TSAL902-L; Sequence=VSP_008635;
CC IsoId=Q9BZP6-2; Sequence=VSP_008635;
CC Name=3; Synonyms=TSAL902-S;
CC IsoId=Q9BZP6-3; Sequence=VSP_008634;
CC -!- TISSUE SPECIFICITY: Highly expressed in stomach. Detected at lower levels in lung.
CC -!- SIMILARITY: Belongs to chitinase class II (family 18 of glycosyl hydrolases).
CC -!- SIMILARITY: Contains 1 chitin-binding type-2 domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; AB025008; BA86980.1; -
CC EMBL; AB025009; BA86981.1; -
CC EMBL; AF290004; AAG60019.1; -
CC EMBL; BC047336; AA47336.1; ALT_INIT.
CC MIM; 606080; -
CC GO; GO:0005737; C:cytoplasm; NAS.
CC GO; GO:0004568; F:chitinase activity; NAS.
CC GO; GO:0006030; F:chitin metabolism; NAS.
CC InterPro; IPR002557; Chitin bind PerA.
CC InterPro; IPR001223; Glyco_hydro_18.
CC InterPro; IPR001579; Glyco_hydro_18AS.
CC Pfam; PF01607; CBM_14; 1.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC SMART; SM00636; Glyco_18; 1.
CC SMART; SM00494; ChitBD2; 1.
CC PROSITE; PS01095; CHITINASE_18; 1.
CC PROSITE; PS01095; CHITINASE_18; 1.
KW Carbohydrate metabolism; Chitin degradation;
KW Polysaccharide degradation; Hydrolase; Glycosidase; Chitin-binding;
```


[illegible]

Signal; Alternative splicing.

CC SIGNAL 21 BY SIMILARITY.

CC CHAIN 22 ACIDIC MAMMALIAN CHITINASE.

CC FT 427 CHITIN-BINDING TYPE-2.

CC FT 426 POLY-SER.

CC FT 420 BY SIMILARITY.

CC FT 140 BY SIMILARITY.

CC FT 51 BY SIMILARITY.

CC FT 372 Missing (in isoform 3).

CC FT 161 /FtId=VSP_008634.

CC FT 108 Missing (in isoform 2).

CC FT 339 /FtId=VSP_008635.

CC FT 432 I -> V (IN REF. 1).

CC FT 432 V -> G (IN REF. 1).

CC SEQUENCE 476 AA; 52271 MW; 92B27BAD2F7EB4CC CRC64;

Query Match 23.5%; Score 708.5; DB 1; Length 476;

Best Local Similarity 35.9%; Pred. No. 7.2e-32;

Matches 149; Conservative 75; Mismatches 134; Indels 37; Gaps 10;

QY 15 RIVCVGTWVSVYHK-VDPYTTIEDIDPFCKTHLMYGFADIKDEYKTYIQVDFPYQDDNHSW 73

DB 23 QLTCTFTWQYRPGGLGRFMPDNIDPCLCTHLIYAFAGRONNEITT-----IEW 71

QY 74 -EKRGYERFNNRLKNPELTTMISLGGWYGESEKYSMAANPTVROQFIQSVDLFQEQYK 132

DB 72 NDVTLYQAEFLGNKNSQLKTLAIGWNGFTGAPFTAMVSTPENRQTFITTSVVKFLRQYE 131

QY 133 FQGLDLDWEYPSGRNLGNPKIDQNYLALVRELKDAFEFHG-----YLLTAAVSPGKDK 185

DB 132 FQGLDLDWEYPSGR-GSPQDQKHLFTVLVQENREAFQEAQKINKPLMTVTAAGISN 190

QY 186 IDRAYDIKELNKLFDMMNMVNTYDYGWENFYGHNAPLYKRPDTEDELHTYFVNNYTHY 245

DB 191 IQSGYEIPQLSQVLDIYHVTVDLHSGWEGYTGENSEPLKYPTDGS-NALNVDVYMWY 249

QY 246 YLNGATRDKLVMGVFPYGRAWSIEDRSKLGLDPAKMGSPPGFISGEGVLSYIELCQL 305

DB 250 WKDNGAPAEKLIIVGPTYGHNFILSNPSNTGIGAPTSGAGPAGPVAKESGIWAYEIC-T 308

QY 306 FQKEWHIQDEYINAPYGNKINWGYDDLASISCKLAFELKELGVSGVMWVSLNDEDFX 365

DB 309 FLKNGATQGDWAPQEPVYAYQGVVWYGYNDINISFDIKAQWLKHNKFGGAMVYALDDDF 368

QY 366 G-HCGP-KNPLLNKVNMMINGDEKNSFECILGPTTPTPTPTPTPTPTPTPTPTPT 418

DB 369 GTFCNQKFFLI-----STLKALGLQASCTAPAQPIETAPSGS 411

RESULT 6

ID CHIT_CABEL STANDARD; PRT; 617 AA.

AC Q11174; O17321;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Probable endochitinase (EC 3.2.1.14).

GN CHT-1 OR C04F6.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Nhan M.;

RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 52-178 FROM N.A.

RX MEDLINE=98324849; PubMed=9662472;

RA de la Vega H., Specht C.A., Liu Y., Robbins P.W.;

RT "Chitinases are a multi-gene family in Aedes, Anopheles and Drosophila".

reduction on day 3 of pregnancy coincident with transport of the embryo from the oviduct to the uterus, a reproductive stage associated with rising progesterone levels.

-!- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U16719; AAC48471.1; -;
ENR1; U17988; AAB01052.1; -;
PIR; I46470; I46470.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 13 13 M -> V (IN REF. 2).
FT CONFLICT 122 122 K -> N (IN REF. 2).
FT CONFLICT 282 282 A -> V (IN REF. 2).
FT CONFLICT 375 375 A -> V (IN REF. 2).
FT CONFLICT 484 484 R -> H (IN REF. 2).
FT CONFLICT 520 520 I -> T (IN REF. 2).
SQ SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;

Query Match 22.6%; Score 682.5; DB 1; Length 539;
Best Local Similarity 34.4%; Pred. No. 2.2e-30;
Matches 171; Conservative 85; Mismatches 180; Indels 61; Gaps 16;

QY 6 HNDYKPMRIVCVGVTWVYHVDPTI--EDIDPEKCTHLMYGFADIDYKTIQVFD 63
DQ 16 HDGGAH--KLVCFPTWNA-FSRGSAIIPRLDPLCLTLVPAFASMNQ--IVPKD 70
QY 64 PYQDHNHNSWKGKRYERFNNLRKNPELTMTISLGGWYEGSEKYSMDAANPTTQROPIQS 123
DQ 71 PLD-----EKILYPERNKLERNRGLTKLLSVGGWNGFTSRFTKMLSTFSNRERFVK 123
QY 124 VLDLQEVKFGDLDDWEYPSRGLGNPKIDKQNYALVRELKDAFEPHG-----YLLT 176
DQ 124 VIALRTHGFDGLDLFFLYPGLR-GSPARDRTWTFVLELLQAFKNEAQLTWPRLLLS 182
QY 177 AAVSGKDKIDRAYDIKELNKLFDWMNVTYDYGWENFYGHNAPLYKRPDDELHTY 236
DQ 183 AAVSGDPHVIQKAYDARLLGLLDFISVLSDYLGWSWEKVTGHSPLFSPLGDPK----- 237
QY 237 FNNVYTHYLLNGATRDKLVMGVPFYGRAWSIEDRSKILGDPAGKMSPPGFTSGBEV 296
DQ 238 -SSAVASYWQLGVPPKELLMGLTYGRTTHLLRASQNELGAGAPASPGYTKQAGE 296
QY 297 LSYIELCOLFQKEHHIYDEYFNAPYNDKIKWYVDLILASICKLAFLKELGVSGWV 356
DQ 297 LAYEVCVSFQAKRWINDQY--VPYAFKGEWYDIDAIISFGYKAPFIKREHFGAMV 354
QY 357 WLENDDPKGH-CGP-KNPLINKVHMMINGDEKSPICILGFTTTPP-----TTTTPP 410
DQ 355 WTLDLDFRGNFGCGTGPFFLAHTLNNLVNDEFSS-----TPSPKFWSTAVNNS 404
QY 411 TTTTTPSTPTTTTTPSTTPP-----TTTTPSTPTTTPSTTPP-TT-----T 455
DQ 405 RTGPEMPTTRDLTGLGILPGGEAVATETHRKSAATTTTPRGETATPTTFLSSGRT 464
QY 456 PAPTSTPTSTPTTHTS 472
DQ 465 AAPEGKTESPGKPLTS 481

RESULT 9
C3L1 MOUSE
ID C3L1 MOUSE STANDARD; PRT; 381 AA.

AC Q61362;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chitinase-3 like protein 1 precursor (Cartilage glycoprotein-39)
DE (GP-39) (BRP39 protein).
GN CH3L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -!- FUNCTION: May play an important role in the capacity of cells to respond to and cope with changes in their environment (By similarity).
CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).
CC -!- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X93035; CAA63603.1; -;
PIR; S61551; S61551.
DR MGD; MGI:1340899; Chi3l1.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 381 CHITINASE-3 LIKE PROTEIN 1.
FT CARBOHYD 60 60 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 381 AA; 43001 MW; EF6581E8184F0450 CRC64;

Query Match 22.6%; Score 680; DB 1; Length 381;
Best Local Similarity 37.1%; Pred. No. 2e-30;
Matches 141; Conservative 75; Mismatches 130; Indels 34; Gaps 10;

QY 15 RIVCYGVTSVYHK-VDPYTIEDIDPFKCTHLMYGFADIDYKTIQVDPDQDNHNSW 73
DQ 23 KLVCFYTSVQYREGVSGSFLDPAIQFLCTHIYSPANIS-----SDNMLSTW 70
QY 74 E---KRGYERFNNLRKNPELTMTISLGGWYEGSEKYSMDAANPTTQROPIQSVDLQOE 130
DQ 71 EWNDESNYDKLNKTKRNTNLKLLSVGGWTKGFKSEIASNTERRTAFVRSVAPFLRS 130
QY 131 YKFDGLDDWEYPSRGLGNPKIDKQNYALVRELKDAF---EP--HGVLTTAAVSPGKD 184
DQ 131 YGFDGLDLAWLYPRLR-----DKQYFSTLILKELNAEFTKEVQPGREKLLLSAALSAKV 184
QY 185 KIDRAYDIKELNKLFDWMNVTYDYGWENFYGHNAPLYKRPDDELHTYFNNVTMH 244
DQ 185 AIDTGYDIAQIAQLDFINLTYDFHGWRFQITGHSPFLFQGGKDT-RPDYNSVNAVQ 243

| | | | |
|----|-----|--|-----|
| QY | 245 | YYLNGATRDKLVGVVFFYGRAWSIEDRSKLGLDPAKMSPPGFTSGEGVLSYIELCQ | 304 |
| DB | 244 | YMIRLGAAQASKLINGITTFGKSPTLAS-SENQLGAPISGEGLPGRFTKEAGTLAYEICD | 302 |
| QY | 305 | LFOKEWHIQDYENYPAGYNKDIXWGYDDLASISKLAFLKELGVSGVMVSWLENDDP | 364 |
| DB | 303 | FLKGAEVHRLSNE--KVPFATKGMMVGVEHKESVKNVKVLTKKELAGAMVALDLDDF | 360 |
| QY | 365 | KGHCGPKN--PLLKNVHMI | 382 |
| DB | 361 | QGTCQPCKEFFPLTNAIKDAL | 380 |

RESULT 10
OGP_MOUSE STANDARD; PRT; 721 AA.

```

AC Q6210;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
GN (Oviductin) (Estrogen-dependent oviduct protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Oviduct;
RC MEDLINE=96115001; PubMed=7492680;
RA Sendai Y.; Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
RA Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:285-294(1995).
CC -! FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
    in the fertilization process and/or early embryonic development.
CC -! SUBCELLULAR LOCATION: Secretory granules.
CC -! TISSUE SPECIFICITY: Epithelial cells of the oviduct.
CC -! SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation
    at the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to licenses@isb-sib.ch).
CC
CC EMBL; D32137; BAA06863.1; -.
CC MGD; MG1:106661; OVGP1.
CC InterPro; IPR001223; Glyco_hydro_18.
CC InterPro; IPR001579; Glyco_hydro_18AS.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC ProDom; PD000471; Glyco_hydro_18; 1.
CC SMART; SM00636; Glyco_18; 1.
CC PROSITE; PS01095; CHITINASE 18; FALSE NEG.
CC KW Glycoprotein; Fertilization; Repeat; Signal.
CC FT SIGNAL 1 21
CC FT CHAIN 22 721
CC FT DOMAIN 486 632
CC FT
CC FT CARBOHYD 402 402
CC FT CARBOHYD 442 442
CC FT CARBOHYD 469 469
CC SEQUENCE 721 AA; 78807 MW; 37246CF8F01665652 CRC64;

```

Query Match Score 680; DB 1; Length 721;
Best Local Similarity 33.7%; Pred. No. 4,1e-30;
Matches 175; Conservative 79; Mismatches 184; Indels 82; Gaps 16

| | | | |
|----|----|--|----|
| QY | 15 | RIVCYGTGSVSVHKVDPTYI--EDIDPFKCHLMYGFAKIDBYKYTIQVDFPYODDNHS | 72 |
|----|----|--|----|

DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT NON_TER 1
FT SIGNAL <1 18
FT CHAIN 19 537 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 537 AA; 59617 MW; CFCDEBF0212D791 CRC64;
Query Match 22.3%; Score 673; DB 1; Length 537;
Best Local Similarity 33.7%; Pred. No. 7.1e-30; Indels 66; Gaps 18;
Matches 170; Conservative 90; Mismatches 179;
QY 6 HNDYSKNPMRIVCVGTWSTVHKVDPTI--EDIDPFKCTHLMYGFADIKDEYKVTIOVFD 63
DB 13 HEDGAH--KLVCYFTNWA-FSPGPGASILPRDLDFLCTHLVAFASNNQ--LVPKD 67
QY 64 PYQDDNHNWSEKRGYERFNRLKNPELTMTISLGWYEGSEKYSMDAANFTYRQOFTOS 123
DB 68 P-QD-----EKILYPEFNKLKERNRELKTLISIGWNFGTSRFTTLMSTFANREKFTAS 123
QY 124 VLDFLOEYKFDGLDLWDWEYPSGRNLGNPKIDKQNYLALVRELKDAPEPHG-----YLLT 176
DB 121 VTALLRTHGFDGLDFLYPGLR-GSPARDRTWTFVLLBELLAQAFKNEAQLTMRPRLLS 179
QY 177 AAVSPGKDKIDRAYDIKELNKLFDWNNVMTYDYGWENFYGHNAPLYKRPDETDELHTY 236
DB 180 AAVSGDPHVQKAYEARLGRLLDFISVLSYDLHGSWEKVTGHSNPLFLSGDPK-----234
QY 237 FNNYTMHYLLNNGATDKLVMPVFPYGRAWSIEDSKLGDPAKMGSPPGFISGEQV 296
DB 235 -SSAYAMNYWRQGVPEKLLMGLPYGRTFHLKASQNELAAQAVGSPGKYTKQAGF 293
QY 297 LSYIELCQLPQKEWHI QVDEYNNAPYGYNDKIKWGYDDLASICKLAFKELGVSQVMV 356
DB 294 LAYEICCFFVRAKRWINDQY--VPYAFKGEWGVYDDAISFGYKAFKIKREHFGAMV 351
QY 357 WSLENDFKGH-CGP-KNPLLNKVMXINGDEKNSEFCLGSPSTTPTP-----TTPTTP 410
DB 352 WTLLDLDFRGYFCGTGPFPLVHTLNLLVNDPSS-----TPSPKFWSTAVNS 401
QY 411 TTPTTPTSPPTTPTTPTPTP-----TTTPTPTPTTPTPTPTPTPTPTPTPTPTPTPT 455
DB 402 RIGPEMPTTRDTLGLILPQGEAVATETHRKSETTTPKGEATPTPLPSGRHT 461
QY 456 PAPTTPTSP-----TTTHTSETP 475
DB 462 AAPEKTESGEXPLTVGHVAVSP 486
ID_OGP_HUMAN STANDARD; PRT; 678 AA.
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Oviduct-specific glycoprotein precursor (Oviductal glycoprotein)
DE (Oviductin) (Estrogen-dependent oviduct protein) (Mucin 9).
DE OVGPI OR OGP OR MUC9.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RA MEDLINE=95119256; PubMed=7819450;
RA Azias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
RT characterization of an estrogen-dependent human oviductal
RT glycoprotein."
RL Biol. Reprod. 51:685-694 (1994).
RN [2]

RP SEQUENCE FROM N.A., AND VARIANT GLN-676.
RA Jaffe R.C.;
RN Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Coville G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: Binds to oocyte zona pellucida in vivo. May play a role
CC in the fertilization process and/or early embryonic development.
CC !- SUBCELLULAR LOCATION: Secretory granules.
CC !- TISSUE SPECIFICITY: Oviduct.
CC !- SIMILARITY: Belongs to family 18 of glycosyl hydrolases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U09550; AAA86946.1; -;
DR EMBL; U58010; AAB04126.1; -;
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.
DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR EMBL; AL390195; CAC36039.1; -;
DR GENE; HGNC:8524; OVGPI.
DR MIM; 603578; -;
DR GO; GO:0007565; P:pregnancy; TAS.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001579; Glyco_hydro_18AS.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR ProDom; PD000471; Glyco_hydro_18; 1.
DR SMART; SM00636; Glyco_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 678
FT CARBOHYD 402 402 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 676 676 E -> Q (in dbSNP:7825).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;
Query Match 22.1%; Score 667.5; DB 1; Length 678;
Best Local Similarity 33.7%; Pred. No. 1.8e-29;
Matches 174; Conservative 86; Mismatches 187; Indels 69; Gaps 18;
QY 6 HNDYSKNPMRIVCVGTWSTVHKVDPTI--EDIDPFKCTHLMYGFADIKDEYKVTIOVFD 63
DB 16 HEDGAH--KLVCYFTNWA-FSPGPGASILPRDLDFLCTHLVAFASNNQI---VAK 69
QY 64 PYQDDNHNWSEKRGYERFNRLKNPELTMTISLGWYEGSEKYSMDAANFTYRQOFTOS 123
DB 70 DLQD-----EKILYPEFNKLKERNRELKTLISIGWNFGTSRFTTLMSTFANREKFTAS 123
QY 124 VLDFLOEYKFDGLDLWDWEYPSGRNLGNPKIDKQNYLALVRELKDAPEPHG-----YLLT 176
DB 124 VISLLRTHGFDGLDFLYPGLR-GSPMHDRTWTFLEIEELLFAFRKDALTMRPRLLS 182

Search completed: March 22, 2004, 06:53:06
Job time : 21.5581 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:40:28 ; Search time 34.8594 Seconds
(without alignments)
1479.047 Million cell updates/sec

Title: US-09-662-293-21
Perfect score: 3014
Sequence: 1 SIKRDNDYSKNPMRIVCVY.....IMPCPGTIWCQKLTIGE 536

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 1007 | 33.4 | 525 | T44445 | chitinase (EC 3.2. |
| 2 | 869 | 28.8 | 1635 | T14075 | chitinase (EC 3.2. |
| 3 | 834 | 27.7 | 554 | A56596 | chitinase (EC 3.2. |
| 4 | 801.5 | 26.6 | 483 | A53918 | chitinase (EC 3.2. |
| 5 | 705.5 | 23.4 | 617 | T15408 | hypothetical prote |
| 6 | 692.5 | 22.6 | 539 | T46470 | estrogen dependent |
| 7 | 673 | 22.3 | 537 | S57197 | oviduct-specific g |
| 8 | 687.5 | 22.1 | 654 | T38605 | oviductal glycopro |
| 9 | 660 | 21.9 | 383 | A49582 | cartilage glycopro |
| 10 | 650 | 21.6 | 383 | S51327 | heparin-binding gl |
| 11 | 646.5 | 21.4 | 405 | S61551 | breast-regressing |
| 12 | 608 | 20.2 | 399 | S27879 | secretory protein |
| 13 | 604.5 | 20.1 | 504 | A38221 | chitinase (EC 3.2. |
| 14 | 588 | 19.5 | 599 | D23784 | chitinase BH0916 [|
| 15 | 561 | 18.6 | 699 | A38368 | chitinase (EC 3.2. |
| 16 | 560 | 18.6 | 1215 | T43916 | chitinase A [impor |
| 17 | 456.5 | 15.1 | 831 | T00323 | chitinase (EC 3.2. |
| 18 | 432.5 | 14.3 | 1484 | T29275 | hypothetical prote |
| 19 | 428 | 14.2 | 869 | T44440 | chitinase (EC 3.2. |
| 20 | 427.5 | 14.2 | 849 | D28510 | chitinase VCA0027 |
| 21 | 424.5 | 14.1 | 2025 | T03884 | hypothetical prote |
| 22 | 414.5 | 13.8 | 563 | S60651 | chitinase precursor |
| 23 | 411 | 13.6 | 423 | J01975 | chitinase (EC 3.2. |
| 24 | 408.5 | 13.6 | 561 | A25090 | chitinase (EC 3.2. |
| 25 | 405.5 | 13.5 | 424 | S68121 | chitinase I precur |
| 26 | 405 | 13.4 | 546 | F84238 | chitinase [impor |
| 27 | 402.5 | 13.4 | 423 | S51369 | chitinase - fungus |
| 28 | 402.5 | 13.4 | 756 | D21452 | chitinase B homolo |
| 29 | 399 | 13.2 | 756 | A81088 | chitinase B homolo |

30 393 13.0 633 2 T24898 hypothetical prote
31 391.5 13.0 558 2 T30418 chitinase (EC 3.2.
32 388 12.9 424 2 S47133 chitinase (EC 3.2.
33 388 12.9 820 2 A40633 chitinase (EC 3.2.
34 385 12.8 799 2 PC4106 chitinase (EC 3.2.
35 381.5 12.7 452 2 JC4038 47K glycoprotein p
36 380 12.6 755 2 T20950 hypothetical prote
37 378 12.5 765 2 T35719 chitinase - Strept
38 377.5 12.5 550 2 T10393 chitinase (EC 3.2.
39 371.5 12.3 398 2 T04761 chitinase homolog
40 368.5 12.2 1088 2 D82246 probable chitinase
41 367.5 12.2 511 2 S61166 probable membrane
42 365 12.1 379 2 T04762 chitinase homolog
43 364.5 12.1 552 2 T41863 chitinase chi-A or
44 360.5 12.0 551 2 G72865 chitinase - Autogr
45 356.5 11.8 427 2 JC4565 chitinase (EC 3.2.

ALIGNMENTS

RESULT 1

T44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C:Species: Anopheles gambiae (African malaria mosquito)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T44445
R:Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z22771
A:Accession: T44445
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-525 <SHE>
A:Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A:Experimental source: adult; gut
C:Genetics:
A:Gene: chi-1
C:Keywords: glycosidase; hydrolase

Query Match 33.4%; Score 1007; DB 2; Length 525;
Best Local Similarity 39.5%; Pred. No. 2.9e-55;
Matches 210; Conservative 86; Mismatches 178; Indels 59; Gaps 12;

QY 6 HNDYKSNPMRIYCYVGTWSVHKVD-PYTIEDIDPFKTHLMYGFADKDEYKYYTQVDFP 64
DB 23 HXAASAEKGVVYVGTWAVYRPGNGRYDIEHIDPSLCTHLMYGFEGINE-DATVRIIDP 81
QY 65 YODDHNINSWEKGYERFNNLRKNBELTMSLGGWYEGSEKYSDMAANPTVYRQOFIOV 124
DB 82 YLDLEEN-WGRGHIKRFVGLKNVPGFKTLAAIGGWNESGRKFSAMAASGEURKRFISDC 140
QY 125 LDPLQYKFDGLDWEYFGSLGNPKIDKQNYLALVRELKDAFPHGYLLTAASVPGKD 184
DB 141 VAPCQRHGFDGLDWEYFPAQDGNPLIDRDHQAQLVEEMREEFHYGLLLTAASVDF 200
QY 185 KIDRAYDIKELNKLFDMMNMVTDYHGGWENFYGHNAPLYKRPDETDELHTYFNNVYTMH 244
DB 201 SAGVSVDIIRSKSHFHLNVWYDMGAWDSYCGINAPLYRGSADTDLRGQINNVASIH 260
QY 245 YLNNAGATDKLVGVVPFYGRAWSIEDRSKLKGPDAKMSPPGFISGEGVLSTELCQ 304
DB 261 FWLAQOCTGRKLVGLPIYGRNFTLASAANTQIGAPTVGCGTVGYRTREPGVNGYNEFCE 320
QY 305 LFQKEWHIQYBYYNAPYGYNDKIVWGYDDIASCKLAFKLKELGVSGVMVWLSLNDDF 364
DB 321 KLATEAWDLRWSEEQVQYAVRNQWGYDDLSRLSVQLKYLDDQLGGAMVSLSTDDF 380
QY 365 KGHC-GPKNPLLNKVNMMINGEKNSFECILGSPSTPTPTPTPTPTPTPTPTPTPTPT 423
DB 381 LGVCGGGRYFLMHEIKSLVNGGT-----PSITTFPFSVAPTT-----STVAGTTT 426
QY 424 TTPSPPT 483

Db 427 TTPGANGTTPPT--SDAPNHTTTTTEGNGPTTRPPSG-----DG 468
Qy 484 HLLKCY--KEGDIHPHNIHKLVC-----EFVNGVWVHIMPGRGTW 526
Db 469 ---PCAGGRYGFVPHPTNCARYICLTADTYEFT-----CPFGTLF 507

RESULT 2
Tl4075
Chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: Tl4075
R:de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.
A:Reference number: 217872
A:Accession: Tl4075
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1635
A:Cross-references: EMBL:AF026492; NID:G2564720; PID:G2564721; PIDN:AA81850.1
C:Genetics:
A:Gene: CHT2
A:introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: Glycosidase; hydrolase; polysaccharide degradation

Query Match 28.8%; Score 869; DB 2; Length 1635;
Best Local Similarity 34.3%; Pred. No. 5e-46;
Matches 196; Conservative 95; Mismatches 196; Indels 84; Gaps 18;

Qy 2 IKRDNYSKNPM-----RIVCVGVTSVHKVD-PYTTIEDIDPFKCTHLMVGFKAIDE 54
Db 650 INKETSNDNAIESVDYKVCYFTWAWYRQNGKYLPEIDADLCTHIVGFAVLR 709
Qy 55 KYTIQVDFPDQDNHNSW---EKRGYERFNRLKNPELTMTLSLGGWEGS-EKYSDM 110
Db 710 DRLVIK-----PHDSWADIDRFRVVEYKKGKKVT--VAIGWNDSDAGDKYSRL 759
Qy 111 AANPYRQGTOSVLDFQEKYDFGLDLDVPGSLGNPK-----IDKQNYLALVRELKD 166
Db 760 VRSAAARQKFIADVAFIEKYGFDGLDLDVFPVCQWDCKGFSDEKEGFASLVVLSQ 819
Qy 167 APEPHGYLLTAASPGKDKIDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPIYKR 226
Db 820 AFKPKGLLSAVSPSKVDEGVDVTLSDYMDWIAVMAYDYGQWDKKTGHVAPYEH 879
Qy 227 PDETDELHYFNVTMNYLNGATRDKLVMGVFPYGRANSIEDRSKLGLDPAKGNP 286
Db 880 PDDFDKT---FNANPTIHYIEKGAIDPRKLVMGMMPYQGSFLADNKEHGLNAKTYGGGE 936
Qy 287 PGFISGEGVLSYIELCOLFQKEWHIQYDEYN-APYGYNDKIWGYDDLASISCKLAF 345
Db 937 AGESTRAGFSYIEICANIRKNTKWTAVDRKGRMGPIAYGQDQVSDDDQYMRHKEY 996
Qy 346 LKELGVSGVWVMSLENDPFKHCQ-PKNPLNKHVNMINGDEKNSFECIL--GPSTTPT 402
Db 997 VKAMELGGAMIVALLDIDFRLNLCDEEYPLLETINVLNYPGPGPCVLEKEQREPR 1056
Qy 403 PTTTPT 462
Db 1057 PTPPT 1104
Qy 463 PSPTTEHTSETPKY-----TTYVDGHLIKYKCEGDIHPHTNIHKLVCPEVNGW 513
Db 1105 PTYTTT---NVFSQEIANEVDEPTCTDGLF-----VHPPTDCKNYICQYK--- 1150
Qy 514 WHINPCPPGTW-----COEKL 532
Db 1151 -----LCPGGLVSWVDHCDWPOSTNCRNKQT 1176

RESULT 3
A56596
Chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
C:Accession: A56596
R:Kramer, K.J.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitin
A:Reference number: A56596; MUID:93357793; PMID:8353525
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:U02270; GB:S64757; NID:G406048; PID:G406049
A:Experimental source: larvae
A:Note: sequence extracted from NCBI backbone (NCBI:136417, NCBI:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 27.7%; Score 834; DB 2; Length 554;
Best Local Similarity 35.9%; Pred. No. 1.8e-44;
Matches 203; Conservative 76; Mismatches 202; Indels 84; Gaps 19;

Qy 10 SKNPRIVCVGVTSVHK-VDPYTIEDIDPFKCTHLMVGFKAIDEYKYTIQVEDPYODD 68
Db 19 SDSRARIVCYFSNWAYTPRGVGYGIEDIPVEKCTHIIYFVIGTEGNSVLIIDPELD- 77
Qy 69 NENSWEKGYERFNRLKNPELTMTLSLGGWEGSEKYSKSHMVAOKSTRMSFIRSVVFL 128
Db 78 ----VDKNFRNFTSLRSHPSVKFVAVGVGWAEGSSKYSHMVAOKSTRMSFIRSVVFL 133
Qy 129 QEVKFDGLDLDWEYPGSR--LGNPKIDKQNYLALVRELKDAF--EPHGYLLTAASPGKDK 185
Db 134 KYDFDGLDLDWEYPGAADRGSGFSKDKFLVLVQELARAFIRVGKGWELTAAPLANFR 193
Qy 186 IDRAYDIKELNKLFDMMNVTYDYGWENFYGHNAPIYKRDEDELTHTYFNVTMNY 245
Db 194 LMEGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSPYKRPD-QWAVEKLVNDGLHL 252
Qy 246 YLNGATRDKLVMGVFPYGRANSIED-----RSKLGLDPAKMSPPGISGEE 294
Db 253 WBEKGCPSNKLVVIGIPFYGRSFTLSAGNNYGLGTFINKEAGGDPAPYTNATGF----- 307
Qy 295 GVLSTYELCOLFQKE--WHIQYDEYNAPYGYNDKIWGYDDLASISCKLAFKELGVS 352
Db 308 --WAYYEICTEVDKDSGWTWKWDEQCKPYAKGTQWGVYEDPRSVELKMWIKQKGYL 365
Qy 353 GVMVWVMSLENDPFKHCQPKNPLNKHVNMINGDEKNSFECILGPSPTTPTPTPTPTPT 411
Db 366 GATWALDMDDFQGLCGEKNPLIKILHKS-----SYVPPPHTEHTTPTP 412
Qy 412 -----TTPPTPS-----PTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 450
Db 413 EMARPPSTSDPSGDPPTTTAKPASTTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 472
Qy 451 P-TTPPTAPT 508
Db 473 PKPEPOPEPEVEVP-ETENE-----VDGSEI-CNSDQDYPDKKHCKDKYWR-- 517
Qy 509 VNGVWVHIMPCTPGTITWCOEKLTC 533
Db 518 VNGE--AMQFSQGHGVFNVELNVC 540

RESULT 4
A53918
Chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C:Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000
C:Accession: A53918
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A:Title: Isolation, cloning, and characterization of new chitinase stored in active form

[illegible]

RESULT 7
S57197
oviduct-specific glycoprotein 95K precursor - bovine (fragment)
C/Species: Bos primigenius taurus (cattle)
C/Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C/Accession: S57197
R/Sendai, Y.; Abe, H.; Kikuchi, M.; Satch, T.; Hoshi, H.
Biol. Reprod. 50: 927-934, 1994
A/Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.
A/Reference number: S57197; MUID:94257768; PMID:8199272
A/Accession: S57197
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-537 <SEN>
A/Cross-references: EMBL:D16639; NID:G9391621; PID:BA04065.1; PID:d1004583; PID:G9391622
C/Keywords: glycoprotein

| | Query March | 22.3%; | Score 673; | DB 2; | Length 537; |
|----|-----------------------|---|--------------------|------------|-------------|
| | Best Local Similarity | 33.7%; | Pred. No. 1.8e-34; | | |
| | Matches 170; | Conservative 90; | Mismatches 179; | Indels 66; | Gaps 18; |
| Qy | 6 | HNYSKNPMRIVCVGTWSVHKVDPTII--EDIDPKFCTHLMYGPFAKIDBYKTKTIQVFD | 63 | | |
| Db | 13 | HHDGAAH--KLIVCYFTNWA-FSRGCPASILPRDLDPFLCTHLLVFAFASMSNQ--IVPKD | 67 | | |
| Qy | 64 | PYQDDNHSWEKGYEFNNLRILKNPELTITISLGGWYEGSEKYSDDMAANPTYRQQTIQS | 123 | | |
| Db | 68 | P-OP-----EKIIYIPRNKIKERNRGKTKLLSIGGNNFGVTIVTLMISTENSRERFVSS | 120 | | |

[illegible]

```
QY 357 WSLENDDPKGH-CGP-KNPLMKVNMINGEKNSEFICLGSTTTPTP-----TTTTPTT 410
      :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 352 WTLDLDDRGYFCGPGPPVLHTNLNVDEFSS-----TPSPKWFSTAVNSS 401
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
QY 411 TTTPTPSPTPTPTTTPTSPPTP-----TTTPSPTTTTTPSPPTPT-TP-----T 455
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db 402 RIGEMPTWTRDLTGGLILPGGEAVATEHRKSETMTIIPKEGIATPTPLTSFGHRT 461
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
QY 456 PAPTTSTPSP-----TTTEHTSERP 475
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db 462 AAPEGTSPGPKPLTTVGHLAVSP 486
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
```

RESULT 8
I38605
oviducta
C/Species
C/Date: 1
C/Access
R/Arias,
Biol. Re
A/Title:
A/Refer
A/Access
A/Status
A/Molecu
A/Residu
A/Cross-

| | | | | | | | |
|-----------------------|--------|---|--------------|------------|------|--------|------|
| Query Match | 22.1%; | Score | 667.5; | DB | 2; | Length | 654; |
| Best Local Similarity | 33.7%; | Pred. | No. 5.2e-34; | | | | |
| Matches | 174; | Conservative | 86; | Mismatches | 187; | Indels | 69; |
| | | | | | | Gaps | 18; |
| QY | 6 | HNDSKNPMELVCVGVGWSVYHVDPVTI--EDIDPEKCTHLMVGFPAKIDEBYKVTIOVFD | 63 | | | | |
| DB | 16 | HHDGAAH--KLVCVFTTWA--HSRPGPASILPHDDUPLFCTHLIFAFASMNNOI---VAK | 69 | | | | |
| QY | 64 | PYODNHNSEWKRGYERFNRLRNKNPBLTTMISLGGWYEGSEKYSMDAANPTYRQOFTQS | 123 | | | | |
| DB | 70 | DLQD-----EKILYPFNKLKERNRELKTLISLGGWNFGTSRTTLMSTFANREKFIAS | 123 | | | | |
| QY | 124 | VLDPLQYKFDGLDWEYFGSRGLGNPKIDKONYALVRELKDAFEPHG-----YLLT | 176 | | | | |
| DB | 124 | VISLRLTHDFDGLDFFLYPLGLR--GSPMHDRWTFLEIEFLFAFRKEALTMFRLLLS | 182 | | | | |
| QY | 177 | AAVSPGKDKIDRAYDIKELANKLFPMXNMVTVDYHGGWENFYGHNAFLYKRPDDELTHTY | 236 | | | | |
| DB | 183 | AAVSGVPHIVOTSVDVRFGLRLDFFINVLSDLLGSWERFTGHRNSPLFSLDEPK----- | 237 | | | | |
| QY | 237 | FNVAHYTHYLLNNGATDKLVGVFPFYGRAWSIEDRSKLLGDPAKMGSPFGFISGEGV | 296 | | | | |
| DB | 238 | --SAYAMNYWRKLAGPSEKLMG:PTVGRTRLLKASKNGIQAARIGAIPASPGYKTKQEGF | 296 | | | | |
| QY | 297 | LSYIELCOLF--QKEBWHIQVDEYNNAPYGVNDKIWGYDDLAISICKLAFELKELGVSGV | 354 | | | | |
| DB | 297 | LAYFEICSFVWGAKKH-----IDYQVVPYANKGKEWGVGDNAISFSYKAWTIRREHFGGA | 352 | | | | |
| QY | 355 | MWLSLENDDFKG--HCGP--KNPLLKNVHMINGDEKNSPECILGSPTTTTPT--PTTPTTP | 410 | | | | |
| DB | 353 | MWTLDMDDVRGTFCGTGPPFLVYLNDILVRAEFS-----STLSPQWLSSAVNSS | 404 | | | | |
| QY | 411 | TTTTPTSPPTPTTTPSPPTTPTTTPSPPTP-----TTTTPTSPPTTPTTPTTPTT | 455 | | | | |
| DB | 405 | STDPERLAVTTAWTTDSKILPPGGEAGVTEHGKCNMTIIPRGITVTPTKETVSLGKHT | 464 | | | | |
| QY | 456 | -----PAPTTSPSPTTTEHTSETP--KYTTVVDGH | 484 | | | | |
| DB | 465 | VALGKTEITGAWTMSVGHOSMTPGEKALTPV--GH | 499 | | | | |

RESULT 9
A49562
cartilage glycoprotein gp39 precursor - human
N:Alternate names: 3K synovial protein
C:Species: Homo sapiens (man)
C>Date: 23-Mar-1995 #sequence revision 23-Mar-
C:Accession: A49562; S16677; A33162
R:Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 269, 25803-25810, 1993
A:Title: Human cartilage gp-39, a major secretory
A:Reference number: A49562; MUID:94064658; PM
A:Accession: A49562
A>Status: preliminary
A:Molecule type: mRNA

```

Db      71  WNDVTLYDTLNTLKNRPENLKTLLSVGNWFGSQRFSKIASNTQSRRTFKSVPPPLRTH 130
Qy      132  KFDGLDLDEWPGSRLGNPKDKQNYLALVRELKDAFEPHG-----YLLTAAVSPGKOK 185
Db      131  GFDGLDLAWISGRR-----DKRHLLTVLKEMCAEFVREALPOTERLLLSGAVSAGKVA 184
Qy      186  IDRAVDIKELNKLFWMNVMYTDYHGGWENFYGHNAPLYKRPDETDELHTYFVNVTMHY 245
Db      185  IDRGYDIAIQISHLDFISLLAYDFHGAWRQTTGHHSPLFRGQGDASS-DRFSNADYAVSY 243
Qy      246  YLNGATRDKLVMGVPFFYGRAWSTEDRSKLLGDPAGKMGPPGFISSGEEGVLSVIELCOL 305
Db      244  VLRLGAPANKLVWGIPPTGRSFTLAS-SKTDVGAPASGPGIPGRFTKKGILAYEICDF 302
Qy      306  FQKEENHIQDEYVYNAPGYNDKLVWGYDDLASISCKLAFKLKELGVSGVMVMSLENDDFK 365
Db      303  LQAT--VRRPLGQQVPVATKGNQVGYDDQESVKNKAKYLKSRQLAGAMVWLLDLDFFR 360
Qy      366  GH-CGS--KNPLLNKNVHNNI 382
Db      361  GNFCGNLRFPLTSAIKDVL 380

RESULT 11
S61551
breast-regressing protein brp39 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S61551; S61550; I48271
R:Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A:Title: neu and ras initiate murine mammary tumors that share genetic markers
A:Reference number: I48271; PMID:95060797; PMID:7970700
A:Accession: S61551
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-405 <MOR>
A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R:Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61550
A:Accession: S61550
A:Molecule type: mRNA
A:Residues: 1-245, 'I', 247-330, 'H', 332-350, 'MVWALDLDLDFQGTQCFKFFPLTNAIKDALA' <MOR>
A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A:Note: the differences at the carboxyl end are due to a frameshift error
C:Genetics:
A:Gene: brp39
C:Superfamily: Streptomyces chitinase chi40
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

Query Match      21.4%; Score 646.5; DB 2; Length 405;
Best Local Similarity 34.4%; Pred. No. 5.6e-33;
Matches 140; Conservative 73; Mismatches 129; Indels 65; Gaps 129;

Qy      15  RIVCYGTWSVYHK-VDPYTTEDIDPFCKTILMYGFAKIDYKTYIQFVPYQDDNHSW 73
Db      23  KLVCYFTWSQYREGVGSGFLDPAIQFFLCTHIIYSFANIS-----SDNNLSTW 70
Qy      74  E---KRGYERFNRLKQPELFTMI:SLGWYEGGEKSDMAANPTYRQCFIQSVLDFLOE 130
Db      71  EWNDESNYDKLNKLKFRNTNLKTLVSQGWKFGKRFSEIASNTERRTAFVRSVAPFLRS 130
Qy      131  YKPDGGLDWEYPSGRNLGNPKDKQNYLALVRELKDAF-----EP--HGYYLLTAAVSPGKD 184
Db      131  YGFDGGLDLAWLYFLRLR-----DKQYFSTLIKELNAEFTKEVQFGREKLLLSAALSAGKV 184
Qy      185  KIDRAYDIKELANKLFDMNMVMTYDHYGGWENFYGHNAPLYKRPDETDELHTYFVNVTMHI 244
Db      185  AIDPYDYDIAIOHLDLFINLMTYDPHGVRROITGHHSPLFGQGDXT-RFRDYSNNVAVO 243

```

QY 245 YYLNNGATRDKLVGMVFFYGRAWSIEDRSKLGDPKAGMSPPGFISGEGVLSYIELCQ 304
Db 244 YNTRLAGAASKLLMGIPTEKGFSTLAS-SENQIGAPISGEGLPGRFTKEAGTAYEICD 302
QY 305 LFQKEEWHIOYDEYNAPGYNDKLVGVDDLASISCKLAFKELGVSGV-----MWSLE 360
Db 303 FLKGAERHLSNE--KVPFATKGNOWGYEDKESVKNYGVFLKELKAGAMCGHWMWI- 359
QY 361 NDDPKGCHCPKPNLLKNVHNMINGDEKNSFCEILGSPSTTTPPTTP 407
Db 360 ----SGHCQPKN-----SSRPTPSRNP 378

RESULT 12

S27879
Secretory protein YM-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999
C:Accession: S27879
R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.
Submitted to the EMBL Data Library, June 1992
A:Description: Molecular characterization of a secretory protein (YM-1) transiently expressed in *Brugia malayi*
A:Reference number: S27879
A:Accession: S27879
A:Molecule type: mRNA
A:Residues: 1-399 <CHA>
A:Cross-references: EMBL:M4584; NID:g202441; PIDN:AB62394.1; PID:g202442
C:Superfamily: Streptomyces chitinase chi40
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-399/Product: secretory protein YM-1 #status predicted <MAT>

Query Match 20.2%; Score 608; DB 2; Length 399;
Best Local Similarity 36.3%; Pred. No. 1.4e-30;
Matches 141; Conservative 67; Mismatches 140; Indels 40; Gaps 14;

QY 15 RIVCVGTWSVYHKVD-PYTIEDIDPFKCTHLMYGAFT--DEYKTYIQVDPYQDDNHN 71
Db 23 QLMCYTWSAKORPIEGSFPGNIDPCLCTHLYAFAGQNEITYT-----H 70
QY 72 SWEKGYERENLRL-LKNPELTMTISLGWYEGSEKYSDMAANPTVROQFIQSVDLFQ 130
Db 71 EQDLRDEALNGLKDKQNTLKTLLAIGWKFGKPAFSAVSTPQNRQFIQSVIRFLAQ 130
QY 131 YKFDGLDWEYPSRIGNPKIKQNYLALVRELKDAFPHG-----YLLTAASVSPK 183
Db 131 YNFDGLNDWQYPSR-GSPDKDHLFSLVKEMKAFEESEVEKDIPRLILT---STGA 186
QY 184 DKTDRAVD-IKELN--KLFPMMNMVYDHYHGGWENFYGHNAPLYKSPDETDLHTYFN 240
Db 187 GIIDVIKSGTRSLNCLSLDIYIQMTYDLDPKDGTYGNSPLYKSPYDGLK-SADLNVD 245
QY 241 YTHYVLLNGATRDKLVGMVFFYGRAWSIEDRSKLGDPKAGMSPPGFISGEGVLSYI 300
Db 246 SIISYWKDNGAASEKLVGPAVGHFTILSDPKTGIGARTISTGPPGKYTDSGLLAY 305
QY 301 ELCOLFQK---EEWHIOYDEYNAPGYNDKLVGVDDLASISCKLAFKELGVSGV 357
Db 306 EVCTFLNEGATEVW---DAPEQVYAYQNEWGYDNRVSFKLKAQWLKDNLLGAVVW 361
QY 358 SLENDDFPKG-HCGPKN-PLLNKVNMIN 383
Db 362 PLDMDFSGSFCHQHFPLTSLKGLN 389

RESULT 13

A38221
Chitinase (EC 3.2.1.14) MF1 - nematode (*Brugia malayi*)
C:Species: *Brugia malayi*
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A38221
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Piessens, W.F.; Perler, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmission-blocking antibodies recognize microfilarial chitinase in *Brugia*

A:Reference number: A38221; MUID:92179220; PMID:1542646

A:Accession: A38221
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <FUH>
A:Cross-references: GB:M73689; NID:g156063; PIDN:AAA27854.1; PID:g156064
A:Note: sequence extracted from NCBI backbone (NCBIP:85345)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 20.1%; Score 604.5; DB 2; Length 504;
Best Local Similarity 30.1%; Pred. No. 3.1e-30;
Matches 156; Conservative 86; Mismatches 208; Indels 69; Gaps 17;

QY 18 CYGTWSVYHKVD-PYTIEDIDPFKCTHLMYGAFTDEYKTYIQVDPYQ-DNHNHSEW 75
Db 27 CYTNWAQYRDGEGKPLGNIPNGLCTHLYAFKVD-----LGDSKPFWEDETEWSK 82
QY 76 RGYERFNILKPELTMTISLGWYEGSEKYSDMAANPTVROQFIQSVDLFQEFKFDG 135
Db 83 GMSAVTKLRETNPGLKVLISYGGYNFGSAIFGTIAKSAOKTERFIKSAFAFKNNFDG 142
QY 136 LLDLWEYP-GSRLGNPKIDKQNYLALVRELKDAFPHGYYLLTAASVSPKDKIDRAYDIKE 194
Db 143 FDLWEIPVGVAEHAKLVAMKTAFFEAKTSQKOR-LLLTAASVAGKGTIDGSYNVES 201
QY 195 LNKLFDMNMVMTYDHYHGGWENFYGHNAPLYKRPDETDLHTYFNVTMYYLNNGATRD 254
Db 202 LGKNFDLLFLMSYDLHGSWEKNVDLHGKLPHTKGEVSGI-GIFNTEPAADYWSKGMPE 260
QY 255 KLVGVVFPYGRAMSIEDRSKLGDPKAGMSPPGFISGEGVLSYIELCQFQKEEWHIQ 314
Db 261 KIITGIPWYAQWTLDPNSETAIGAAASRPSSASKTNPAGGTASWEICKYKEGGKETV 320
QY 315 YDEYNAPGYNDKLVGVDDLASISCKLAFKELGVSGVWVMSLENDDFKG-HCGP-KN 372
Db 321 HQEGVGA-YMVKGQWYDNEETIRIKMKWLKKGYGGAFIWALDDDFDTGKSCGKGPY 379
QY 373 PLLNKHVMINGDEKNSFCEILGSPSTTTPPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 432
Db 380 PLLNAISSELGESN-----PEITTEPSITETAYETDEE----- 418
QY 433 TTPSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 488
Db 419 -----TSETEAYDTD-----ETESET-EATTYDITDETEGQ-EC 451
QY 489 -YKEDIHPNINIKYLVCEFWNGWVWVHMPCPPTGIIW 526
Db 452 PERDGLPPTDCHLFIOQ-----ANNIAYVMQCPATTFF 486

RESULT 14

D83764
Chitinase BH0916 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83764
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: D83764
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04635.1; GSPDB:GN001
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0916

Query Match 19.5%; Score 588; DB 2; Length 599;
Best Local Similarity 29.6%; Pred. No. 4.1e-29;
Matches 179; Conservative 91; Mismatches 200; Indels 134; Gaps 29;

```
QY 10 SKNPMRIYCVGTWTSVHKVDYPTIEDIDPKCTHLMYGFAC----- 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 36 SDDQKIVAYPSWGAVER--DYQWDDIDASKISHINYAFANICWDGHHGNDPAGPNPQ 93
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 52 -----IDYKYTIQVDPYOD-----DNHNSWEK--RG-YERFNNLRLKNPELTTM 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 TWSQDENGVIDVNGSVMDPMDIDAGKSPGDTWDEPLRGNFKQLNKLKEEHPHLKTL 153
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 ISLGGWEGSEKYSMAANPYRQOPIOSVLDLQCYKFGDLDDWEYP--GSLGNPK- 151
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 ISVGGW-TWSNRFSDMAATKTRENFANSVAFIRKYGFDGVDVDWEYPVSGGPGNSRR 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 -IDKQNLALVRELKDAPEPHG-----YLLTAA--VSPGKDKIDRAYDIKELNKLFDWM 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 PEDKENHVLVQLQEVDRKLDDEAGQDCKDYLLTIASGASPGYVNNK---LNEIAEIVDWI 269
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 NMVTDYHGWNFYGHNAPLYKRPDET-ELHT--YFNVNTMYYLNNNGATRDKLVMG 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 NIMTYDFNGMCNIGSHNAPLYDPATANTELPTPEHFNVESAVEGHLQAGVPEHKLVLG 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 VPFGRAWISIEDRSKLGDPAGKMPGPFISGEGVLSYIELCOLFOKEBWHIQ---YD 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 330 MPFYGRGWSNCDGA--NQGEYQR-CAPPREGTWENGVEFDSDL-----EDHYINQNGYQ 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 317 EYTN---APYGYN--DKIWWGYDDLASICKLAPLKLKELGVSGVMVMSLENDDFKHGCP 370
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 381 RYWDVAVKVPFLYNATNGNFITYDDEESFRYKTDIFKSNLNGSMFWDVSGDR----- 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 371 KNPLLNKVNMMINGDEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 430
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 434 -----NGILLTALADQLG---FTHEGQEPPEPSAPNIOATEVTSITVLT 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 431 ---PTTPTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 WOAPTEPTQVSVAYDSKEKTHITITIEDLQPETTYTFVVSAEHKDGIHAGQALQVT 538
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 475 PKITYVDG-----HLIKCYKEG-IEHPNTNHLKLVCEFVNGGWNWHIMPCPPGT--- 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 TKSETGGDGTAPTQWANNVYTGQVQVHGKGLYE-----AKWMT--TCEBPGTGTGE 588
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 525 --IW 526
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 589 MGWV 592
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
A38368
Chitinase (EC 3.2.1.14) precursor - Bacillus circulans
C:Species: Bacillus circulans
C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 15-Oct-1999
C:Accession: A38368
R:Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.
J. Biol. Chem. 265, 15659-15665, 1990
A:Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolution
A:Reference number: A38368; MUID:90368776; PMID:2203782
A:Accession: A38368
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-699 <WAT>
A:Cross-references: GB:M57601; GB:J05599; NID:g1066341; PIDN:AAA81528.1; PID:g142688
C:Superfamily: fibronectin type III repeat homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 18.6%; Score 561; DB 2; Length 699;
Best Local Similarity 29.4%; Pred. NO. 2.4e-27;
Matches 170; Conservative 67; Mismatches 185; Indels 156; Gaps 23;

QY 15 RIVCYGTWTSVHKVDYPTIEDIDPKCTHLMYGFACI----- 52
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 KIVGYTPSAAAYGR--NYNVADIDPTKVTHINYAFADICWNGIHGNPDPSGPNVTTWCO 102
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 DEYKYTIQV-----FDPYQDDNH---NSWEK---RGYERFNNLRLKNPELTTMISLG 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: March 22, 2004, 07:01:31
Job time : 35.8594 secs

```
Db 103 NEKSOTINVPNGTIVLGDWIDTGTFTAGDTWDQPIAGNINOLNKLKQTNPNLKTIIISVG 162
QY 99 GWYEGSEKYSMAANPYRQOFTQSVDLQCYKFGDLDDWEYP--GSLGNPK--IDK 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 GW-TWNRFSVDAATATREVFANSVDFLRKYNFDGVDLWEYPVSGGLDGNKPERDK 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 QNYLALVRELKDAPEPHG-----YLLTAAVSPGKDKIDRAYDIKELNKLFDWMNVTYD 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 222 QNYTLLSKIREKLDAGAVDGKKYLLTTIA--SGASATYAANTELAIAAIVDWINIMTYD 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 209 YHGGWENFYGHNAPLYKRPDET-----DELHTYFNVNTMYYLNNNGATRDKLVMGVPEYG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 FNGAWOKISAHNAPLNYDPAASAGVPDANT-FNVAAGAQGHLDAGVPAKLVLGVPFFYG 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 RANSIEDRSKLGDPAGKMPGPFISGEGVLSYIELCOLFOKEBWHIQ---YDEYYN- 320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 RGWDGCAQAGNGOYQCTCGSSVG--TWAGSDFDYDL-----EANYLNKNGYTRYWND 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 ---APYGYN--DKIWWGYDDLASICKLAPLKLKELGVSGVMVMSLENDDFKHGCPKNPLL 375
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 392 TAKVPVLYNASNKRFSYDDAESVGYKTAIKSKGLGAMFELSGDRNK----- 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 376 NKVHNMINGDEKNSFECILGPSTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 -TLQNLKAD-----LPTGTVPPVDITAPSPVGNARSTGVTANSVTLAWN 486
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 418 -----SPTTPTTPTTPTTPTT-----PSPPT 439
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 ASTDNVGVTCGYNYNGANLATSVTGTATISGLTAGTSYFTTIKAKDAAGNLSAASNAVT 546
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 440 PTTTPTSP--TTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 474
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 547 VSTTAQPGGDTQAPTAPTNLASTAQTSSITLTSWTAST 584
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2004, 06:42:54 ; Search time 43.2018 Seconds
(without alignments)
640.518 Million cell updates/sec

Title: US-09-662-293-21
Perfect score: 3014
Sequence: 1 SIKRDNDYSKNPVRIVCV.....IMPCPPGTWQCKLTGICE 536
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
6: /cgn2_6/prodata/2/iaa/5B COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------|-------------------|
| 1 | 3014 | 100.0 | 536 | US-09-292-225-21 | Sequence 21, Appl |
| 2 | 3014 | 100.0 | 555 | US-09-292-225-15 | Sequence 15, Appl |
| 3 | 3014 | 100.0 | 555 | US-09-292-225-18 | Sequence 18, Appl |
| 4 | 2475 | 82.1 | 490 | US-09-292-225-41 | Sequence 41, Appl |
| 5 | 2475 | 82.1 | 509 | US-09-292-225-35 | Sequence 35, Appl |
| 6 | 2475 | 82.1 | 509 | US-09-292-225-38 | Sequence 38, Appl |
| 7 | 834 | 27.7 | 554 | US-08-524-051-2 | Sequence 2, Appli |
| 8 | 834 | 27.7 | 554 | US-09-052-778-16 | Sequence 16, Appl |
| 9 | 812.5 | 27.0 | 583 | US-09-545-814-2 | Sequence 2, Appli |
| 10 | 812.5 | 27.0 | 583 | US-09-545-814-5 | Sequence 5, Appli |
| 11 | 811.5 | 26.9 | 559 | US-09-545-814-14 | Sequence 14, Appl |
| 12 | 809.5 | 26.9 | 635 | US-09-545-814-32 | Sequence 32, Appl |
| 13 | 738 | 24.5 | 466 | US-08-486-839-4 | Sequence 4, Appli |
| 14 | 738 | 24.5 | 466 | US-09-151-011-4 | Sequence 4, Appli |
| 15 | 738 | 24.5 | 466 | US-09-039-198A-2 | Sequence 2, Appli |
| 16 | 738 | 24.5 | 466 | US-09-343-623-4 | Sequence 4, Appli |
| 17 | 738 | 24.5 | 466 | US-08-877-599-2 | Sequence 2, Appli |
| 18 | 738 | 24.5 | 466 | US-09-367-574-2 | Sequence 2, Appli |
| 19 | 732.5 | 24.3 | 387 | US-08-486-839-6 | Sequence 6, Appli |
| 20 | 732.5 | 24.3 | 387 | US-09-151-011-6 | Sequence 6, Appli |
| 21 | 732.5 | 24.3 | 387 | US-09-343-623-6 | Sequence 6, Appli |
| 22 | 732 | 24.3 | 466 | US-09-039-198A-4 | Sequence 4, Appli |
| 23 | 732 | 24.3 | 466 | US-08-877-599-4 | Sequence 4, Appli |
| 24 | 732 | 24.3 | 466 | US-09-267-574-4 | Sequence 4, Appli |
| 25 | 721.5 | 23.9 | 373 | US-09-039-198A-14 | Sequence 14, Appl |
| 26 | 721.5 | 23.9 | 373 | US-09-039-198A-15 | Sequence 15, Appl |
| 27 | 721.5 | 23.9 | 373 | US-08-877-599-14 | Sequence 14, Appl |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 28 | 721.5 | 23.9 | 373 | 4 | US-08-877-599-15 | Sequence 15, Appl |
| 29 | 721.5 | 23.9 | 373 | 4 | US-09-267-574-14 | Sequence 14, Appl |
| 30 | 721.5 | 23.9 | 373 | 4 | US-09-267-574-15 | Sequence 15, Appl |
| 31 | 672.5 | 22.3 | 489 | 4 | US-09-545-814-29 | Sequence 29, Appl |
| 32 | 650 | 21.6 | 383 | 4 | US-09-459-749D-17 | Sequence 17, Appl |
| 33 | 630 | 20.9 | 385 | 2 | US-08-694-915-2 | Sequence 2, Appli |
| 34 | 630 | 20.9 | 416 | 2 | US-08-694-915-4 | Sequence 4, Appli |
| 35 | 630 | 20.9 | 423 | 4 | US-08-850-348A-2 | Sequence 2, Appli |
| 36 | 419.5 | 13.9 | 700 | 4 | US-09-408-647A-2 | Sequence 2, Appli |
| 37 | 415 | 13.8 | 440 | 3 | US-09-052-778-13 | Sequence 13, Appl |
| 38 | 411.5 | 13.7 | 442 | 3 | US-09-052-778-2 | Sequence 2, Appli |
| 39 | 411 | 13.6 | 389 | 1 | US-07-939-501A-1 | Sequence 1, Appli |
| 40 | 411 | 13.6 | 389 | 3 | US-08-448-398-7 | Sequence 7, Appli |
| 41 | 411 | 13.6 | 423 | 1 | US-07-939-501A-10 | Sequence 10, Appl |
| 42 | 411 | 13.6 | 423 | 1 | US-07-939-501A-12 | Sequence 12, Appl |
| 43 | 403.5 | 13.4 | 561 | 1 | US-08-358-901-2 | Sequence 2, Appli |
| 44 | 403.5 | 13.4 | 561 | 1 | US-08-566-347-2 | Sequence 2, Appli |
| 45 | 403.5 | 13.4 | 561 | 1 | US-08-693-835-2 | Sequence 2, Appli |

ALIGNMENTS

RESULT 1
US-09-292-225-21
; Sequence 21, Application US/09292225
; Patent No. 6455586
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-21

| | | | | |
|-----------------------|--------------|--|---------------|-------------|
| Query Match | 100.0% | Score 3014; | DB 4; | Length 536; |
| Best Local Similarity | 100.0% | Pred. No. 1.2e-232; | Mismatches 0; | Indels 0; |
| Matches 536; | Conservative | 0; | | Gaps 0; |
| QY | 1 | SIKRDNDYSKNPVRIVCVGTWSYHKVDPTTIEDIDPFKTHLMYGFAXIDYKTYIQ | 60 | |
| Db | 1 | SIKRDNDYSKNPVRIVCVGTWSYHKVDPTTIEDIDPFKTHLMYGFAXIDYKTYIQ | 60 | |
| QY | 61 | VDPFDQDDNHNHSEKRGYERFNRLKNPDLTWTISLGGWYEGSEKYSMDAANPTYROOF | 120 | |
| Db | 61 | VDPFDQDDNHNHSEKRGYERFNRLKNPDLTWTISLGGWYEGSEKYSMDAANPTYROOF | 120 | |
| QY | 121 | TOSVLDFLOEYKFDGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHGYLLTAUVS | 180 | |
| Db | 121 | TOSVLDFLOEYKFDGLDLDWEYVGSRLGNPKIDKQNYLALVRELKDAPEPHGYLLTAUVS | 180 | |
| QY | 181 | PGKDKIDRAYDIKELNKLFDWNNVNTYDHGWFENFYGHNAFLYKRPDETDELHTYFNVN | 240 | |
| Db | 181 | PGKDKIDRAYDIKELNKLFDWNNVNTYDHGWFENFYGHNAFLYKRPDETDELHTYFNVN | 240 | |
| QY | 241 | YTMHYLLNNGATRDKLVMGVFPFYGRAWSIEDRSKLKLGDPKGMSPPGFISGEEGVLSYI | 300 | |

Db 241 YTHYLLNNGATRDKLVGMVGFYFGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYI 300
Qy 301 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 360
Db 301 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 360
Qy 361 NDDFKHCGPKNPLNKNVHNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPT 420
Db 361 NDDFKHCGPKNPLNKNVHNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPT 420
Qy 421 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
Db 421 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
Qy 481 VDGLHIKCYKEGDIHPHTNIHKYLVCDFVNGVWVHIMPCPPGTIWCQEKLTICGE 536
Db 481 VDGLHIKCYKEGDIHPHTNIHKYLVCDFVNGVWVHIMPCPPGTIWCQEKLTICGE 536

RESULT 2
US-09-292-225-15
; Sequence 15, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-15

Query Match 100.0%; Score 3014; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.3e-232;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SIKRDNDYSKNPVRIVCYVGTWSVYHKVDPTYTIEDIDPFKTHLMYGAKEIDYKTYIQ 60
Db 20 SIKRDNDYSKNPVRIVCYVGTWSVYHKVDPTYTIEDIDPFKTHLMYGAKEIDYKTYIQ 79
Qy 61 VFDPYQDDNHSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMAANPTYRQOF 120
Db 80 VFDPYQDDNHSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMAANPTYRQOF 139
Qy 121 IQSVLDFLOEYKFDGLDWEYPGSLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAVS 180
Db 140 IQSVLDFLOEYKFDGLDWEYPGSLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAVS 199
Qy 181 PGKDKIDRAYDIDKELNKLFDWMNVMTYDHGWNFYGHNAPLYKRPDDETLHTYFNVN 240
Db 200 PGKDKIDRAYDIDKELNKLFDWMNVMTYDHGWNFYGHNAPLYKRPDDETLHTYFNVN 259
Qy 241 YTHYLLNNGATRDKLVGMVGFYFGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYI 300
Db 260 YTHYLLNNGATRDKLVGMVGFYFGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYI 319

Qy 301 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 360
Db 320 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 379
Qy 361 NDDFKHCGPKNPLNKNVHNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPT 420
Db 380 NDDFKHCGPKNPLNKNVHNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPT 439
Qy 421 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 480
Db 440 TPTTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 499
Qy 481 VDGLHIKCYKEGDIHPHTNIHKYLVCDFVNGVWVHIMPCPPGTIWCQEKLTICGE 536
Db 500 VDGLHIKCYKEGDIHPHTNIHKYLVCDFVNGVWVHIMPCPPGTIWCQEKLTICGE 555

RESULT 3
US-09-292-225-18
; Sequence 18, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-18

Query Match 100.0%; Score 3014; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.3e-232;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SIKRDNDYSKNPVRIVCYVGTWSVYHKVDPTYTIEDIDPFKTHLMYGAKEIDYKTYIQ 60
Db 20 SIKRDNDYSKNPVRIVCYVGTWSVYHKVDPTYTIEDIDPFKTHLMYGAKEIDYKTYIQ 79
Qy 61 VFDPYQDDNHSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMAANPTYRQOF 120
Db 80 VFDPYQDDNHSWEKRGYERFNNLRKNPELTMTISLGWYEGSEKYSMAANPTYRQOF 139
Qy 121 IQSVLDFLOEYKFDGLDWEYPGSLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAVS 180
Db 140 IQSVLDFLOEYKFDGLDWEYPGSLGNPKIDKQNYLALVRELKDAFEPHGYLLTAAVS 199
Qy 181 PGKDKIDRAYDIDKELNKLFDWMNVMTYDHGWNFYGHNAPLYKRPDDETLHTYFNVN 240
Db 200 PGKDKIDRAYDIDKELNKLFDWMNVMTYDHGWNFYGHNAPLYKRPDDETLHTYFNVN 259
Qy 241 YTHYLLNNGATRDKLVGMVGFYFGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYI 300
Db 260 YTHYLLNNGATRDKLVGMVGFYFGRAWSIEDRSKLGDPAGKMSPPGFISGEGVLSYI 319
Qy 301 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 360
Db 320 ELQCFQKEEWHIOYDEYNAPYNDKIWGYDDLASISCKLAFKELGVSQVWVMSLE 379


```

295 GVLSYIELCOLFOKEE--WHIQYDEYINAPGYNDKIWVGDDLASISCKLAFKLKELGVS 352
308 --WAYYEICTEYVDKDSGWTKKWDQGGKCPYAYKGTQWGVYDFRSVEIKMNIKOKGYL 365
353 GVMVWSLENDDFKGHGCGKPNLLKNVKNMINGDEKNSFECILGSPSTTPTPTPTPTPT- 411
366 GAWTWAIDMDDFGLCGCKNPKLIKILHKHMS-----SYVPPPHHTENTTPTPT 412
412 -----TPTPTPS-----PTPTPTTPTPTTPTPTTPTPTTPTTPTT-----PT 450
413 EWARPPSTSDSEGDPIPTTTAKPASTTKTTVKTTTTTAKPPQSVIDEENDINVRPE 472
451 P-TPTTAPPTTSPPTTTEHTSETPKVTYVDGHLIKYCKEKG-IPHPNTNHHKYLVCF 508
473 PKPEPQPEVEVP-PTENE-----VDGSEI-CNSDQDYIPDKKHCDKYWRC-- 517
509 VNGGWWVHIMPCTPGTINCOEKLTC 533
518 VNGE--AMQFSQHGTVFVNLNVC 540

RESULT 8
US-09-052-778-16
; Sequence 16, Application US/09052778A
; Patent No. 6060590
; GENERAL INFORMATION:
; APPLICANT: Bryant, Peter J.
; APPLICANT: Kawamura, Kazuo
; TITLE OF INVENTION: CHITINASE RELATED PROTEINS AND METHODS
; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 07306/015001
; CURRENT APPLICATION NUMBER: US/09/052,778A
; CURRENT FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Manduca sexta
US-09-052-778-16

Query Match 27.7%; Score 834; DB 3; Length 554;
Best Local Similarity 35.9%; Pred. No. 1.8e-58;
Matches 203; Conservative 76; Mismatches 202; Indels 84; Gaps 19

QY 10 SKNPMEIVCVGWTGSVYHK-VDPYTIEDIDPFKCTHLMYGPAKIDEXKYTIQVDFPYQDD 68
DB 19 SDSRAIVCYFNSWAVYRPGVGYGIEDIPVEKCTHIISYFIVTEGNSVLIIDPELD- 77
QY 69 NHNSWEKRGYERFNRLRLKNLPETTWISLGWYEGSEKYSDMAANPTYRQOFIQSVLDL 128
DB 78 ----VDKNGFRNFTSLRSSHPSVKPFMAVGVGAEGSSKYSHMVAQKSTRSFIRSVVSFL 133
QY 129 QEYKFGDLDLWEYPSGR-LGNPKTDKNYLAALVRELKDAF--BPHGYLLTAAVSPGOK 185
DB 134 KKYDFDGLDLWEYPGAADRGGSFSDKDKFLYLVQELARAFIRYKGMWELTAAPLANFR 193
QY 186 IDRAYIDIKELNKLFDWMNTYTDYHGWFENFYGHNAPLYKEPDETDELHYFYVWNTMY 245
DB 194 IMEGYHVPELCOELDAIHVMSYDLRGNWAGFADVHSPLYKRPD-QWAYEKLNVNDGLHL 252
QY 246 YLNGATRDCLKVNGVPFYGRAWSIED-----RSKLKLGDPAGKMSPPGFISGEE 294
DB 253 WESKGCPSNKLVGVIFFYGRSFTLSAGNNNYGLGTFINKEAGGDPAPYTNATGF----- 307
QY 295 GVLSYIELCOLFOKEE--WHIQYDEYINAPGYNDKIWVGDDLASISCKLAFKLKELGVS 352
308 --WAYYEICTEYVDKDSGWTKKWDQGGKCPYAYKGTQWGVYDFRSVEIKMNIKOKGYL 365
QY 353 GVMVWSLENDDFKGHGCGKPNLLKNVKNMINGDEKNSFECILGSPSTTPTPTPTPTPT- 411
DB 366 GAWTWAIDMDDFGLCGCKNPKLIKILHKHMS-----SYVPPPHHTENTTPTPT 412

```

Dd 478 HEEKSEQDNQVGSQDATTATD---VDCSQSDYLPH-EDCNKYIRC-----VHGEAV 525

Qy 517 IMPCPPGTIW 526
: : ||::

Dd 526 LFTCREGTVY 535

RESULT 10
US-09-545-814-5
; Sequence 5, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-5-CI
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-5

Query Match 27.0%; Score 812.5; DB 4; Length 583;
Best Local Similarity 35.6%; Pred. No. 1e-56;
Matches 196; Conservative 86; Mismatches 197; Indels 71; Gaps 22

Qy 10 SKNPMRIVCVGTWSVYHK-VDPYTIEDIDPFKCTHLMYGPAKIDEVKYTIQVDFPDQDD 68
: ||||| :
Db 24 SDQKARIVCFYNWAVYRGIRYGIEDIPVDLCTHIIVSIGVDDDKDMSVLVIDPELDI 83
: ||||| :
Qy 69 NHNSWEKRGYERFNLRNLKANPELLTMI SLIGWYEGSESKYSDMAANPYRQQFIQSVDLF 128
: ||||| :
Db 84 DDN----GPKFTNLRIKHPNVKLQIAVGGAEGGKKYSTWAEEKRSAFIRSVDPM 138
: ||||| :
Qy 129 QEYKFGLDLWEYPGSR-LGNPKIKIDKNYLAALVRELKDAPFPG--YLTAAVSPGOK 185
: ||||| :
Db 139 NEYKFGFDLWEYFGAADRGGSFSDKFKFLYFVQELRRAPFNKQKNWEITMAVPIAKFR 198
: ||||| :
Qy 186 IDEAYDIKELNKLFWMMVMTYDYHGWHENFVGHNAPLYKRPDETDELHTYFVNVTMHY 245
: ||||| :
Db 199 LOEYHVPELCESLDAIHVMSYDLRGNWAGADTHSLPYRRPHD-QVAYEKLVNDGLQL 257
: ||||| :
Qy 246 YLNGNATRDKLVMPGFYGRAWSIEDRK-LKLG---DPAKGMSPPGFIAGEGVLSYE 301
: ||||| :
Db 258 WVDMGCPANKLVGVFPFYGRSTLNSNKDYELGTYNKEAGGSGEPGYTNATGFSIYE 317
: ||||| :
Qy 302 LCOLFQ--KEEWHIQDEYYNAPGYNDKIWGVDYDLASISCKLAFUKELGVSQVMWSL 359
: ||||| :
Db 318 ICLEVDPDGKWTCKWDEHGKVPYAYKGNQWGVYEDPKSVALKMEFIKSKGYGGAMTWAI 377
: ||||| :
Qy 360 ENDDFFKGHC-PKNPLLNKVHNMINGDENKSFCILGPSTTTTP--TTTTPTPTTTPTT 416
: ||||| :
Db 378 DMDDFQGVCSDDKHTLATIMHDYM---KNVTFEFDSSRITPRFEWAKPSTPSQEP-D 432
: ||||| :
Qy 417 PSPTTPTT-TBSPST-TPTTTPSTPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469
: ||||| :
Db 433 DTPIYITTHAPRSRKTTRKPPT--TTTVAAATTPVAT-----TTTEHHHH 477
: ||||| :
Qy 470 HTSETP-----KYTTYVDGHLIKCYKEGDIPHPTNIHKYLCBFVNGWMVH---- 516
: ||||| :
Db 478 HEEKSEQDNQVGSQDATTATD---VDCSQSDYLPH-EDCNKYIRC-----VHGEAV 525

Qy 517 IMPCPPGTIW 526
: : ||::

Dd 526 LFTCREGTVY 535

```

RESULT 11
US-09-545-814-14
; Sequence 14, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USSES THEROSF
; FILE REFERENCES: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-14

Query Match          26.9%; Score 811.5; DB 4; Length 559;
Best Local Similarity 35.8%; Pred. No. 1.1e-56;
Matches 195; Conservative 86; Mismatches 193; Indels 71; Gaps 22;

Qy      15 RIVCYGTGWSYHK-VDPYTIETIDPKCTHLMYGFAKIDKYKTIQVDFYQDDNNHSW 73
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      5 RIVCFPSNWAVYPGIGRYGIEDIPVDLCTHIVSFIGVDDKOWSLVIDEFLDIDN-- 62

Qy      74 EKGYERFNNLRINKPELTMTISLGGWYGESEKYSDMAANTPYEQQTQSVLDFLQEKYF 133
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      63 ---GFKNFTNLRKTHPNVKLIQIAGVGAEGGKKYSTWAEKRSKSAFIRSVVDPMWYKF 119

Qy      134 DGLDLDWEYEGSR-LGNPKDKQNYLALVELKDAPPHG--YLLTAIVSPGKKIDRAY 190
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      120 DGFDLDEYFGAADRGGSGFSDKOKFLYFVQELRFAFNKQGNWEITWAVPIAKPRLQEGY 179

Qy      191 DIKELNKLFDWMNVMTVDYHGGWENFYGHNAPLYKRPDETDLHTYFNVTMTWYVYLNNG 250
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      180 HVPCLCELLDAIHVMSYDLRGNWAGFADTHSPLYRRPHD-QYAYEKUNVDGLQIWDVG 238

Qy      251 ATRDKLVMGVPFYGRAWSIEDRSK-LKLG---DPAKGMSPFGISGEEGLSYIELCOLF 306
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      239 CPANKLVGVPPFYGRSFTLSNSKDKYRLGTVINKEAGGEGPGPYTNATGFSYVEICLEV 298

Qy      307 Q--KEEWHIQVDEYNAPYGNDKIWGVYDDLASISKPLAPLKELGSGVMVWLENDDF 364
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      299 DDPESKGTWKWDEHGKYPVAYKGNQWGVYEDPKSVALKMEFIKSGYCGGAMTWAIDMDF 358

Qy      365 KGHCG-PKNPLNKKVHNMINGDENKSECIILGPSTTPTP--TTTTPTPTTPTTPTSPPT 421
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      359 QGVCSDDKHTLAVIMHDYM---KNYIVPFDSSRITPRPWAKPSPSPQEP-DDTPIYI 413

Qy      422 PTT-TPSPT-TPTTPTSPPTTPTTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 474
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      414 PTHAPPSRKPTRKPKPT--TTTVAATTPVAT-----TTTEHHHHHHHEEK 458

Qy      475 P-----KYTTVVBGHLIKCYKEGDIPHTNIHKYLVCFYNGGWVH-----IMCP 521
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      459 PSEQDNQVGSQDTTATD---VDCSQEDYLPH-EDCNKYRC-----VHGEAVLFCR 506

Qy      522 PGTIW 526
      |||:
Db      507 EGIVY 511
      |||:

```

RESULT 12
US-09-545-814-32
; Sequence 32, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Bechter, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND

```

; TITLE OF INVENTION:  US09-662-293-21
; FILE REFERENCE:  FC-5-CI
; CURRENT APPLICATION NUMBER:  US/09/545,814
; CURRENT FILING DATE:  2000-04-07
; PRIOR APPLICATION NUMBER:  60/128,833
; PRIOR FILING DATE:  1999-04-09
; NUMBER OF SEQ ID NOS:  40
; SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH:  635
; TYPE:  PRT
; ORGANISM:  Tagged Ctenocephalides felis
; US-09-545-814-32

Query Match      26.9%; Score 809.5; DB 4; Length 635;
Best Local Similarity 35.6%; Pred. No. 2e-56;
Matches 196; Conservative 85; Mismatches 198; Indels 71; Gaps 22;

QY 10 SKNPMRIVCVGWSVYHK-VDPYTTIEDDPFKCTHMYGFAKIDKYKTYIQVDFPQYD 68
Db 25 SDQKARIVCVFNSWAVYRPGICRYGIEDIPVDLCTHIVSFVIGVDDKDSVLDVDPDL 84
QY 69 NNSWKRGRYERFNNLRKNPELTWISLGGWYEGSEKYSDMAANPTYRQFIQSVDPL 128
Db 85 DDN-----GPKNTNLRKTHPNVKQLAVGWAEGGKYSTWVAEKRSAPFIRSVVDFM 139
QY 129 QBYKFDGLDLDWYPSGR-LGNPKIDKQNYLALVRELKDAFEPHG--YLLTAAVSPGDK 185
Db 140 NEYKFDGLDLDWYPSGR-LGNPKIDKQNYLALVRELKDAFEPHG--YLLTAAVSPGDK 199
QY 186 IDRAYDIKELNKLFDWNNVMTYDVGWENFYGHNAFLYKRPDETDELHTYFNVTMHY 245
Db 200 LQGYHVPELCELULDAIHVMSYDLRGNWAGFATHSPLYRRPHD-QYAYEKLNVNDGLQL 258
QY 246 YLANGATRDKLVMGVFPYGRAMWEDRSK-LKLQ---DPAKGMSPGPGFISGEEGLSYIE 301
Db 259 WVDNGCPANKLVGVFPYGRFPTLSNKKDYLRTYINKEAGGEGPGYTNATGFSIYIE 318
QY 302 LCQLFQ--KEEWHIOYDEYNNAPYGYNDKTIWGYDDLASISCKLAFKELGVSGVMVWSL 359
Db 319 ICLEVDPSKGTWKWDEHKGVYAYKGNQWGVYEDPKSVALKWFIKSGYGGAMTWA 378
QY 360 ENDDFKHCG-PKNPLLNKVNHNMGDEKNSFCILGPSTTTTP--TTTPTTPTTPTT 416
Db 379 DMDDFQVCSDRDKHTLAVIMHDYN---KNYIVPEFDSRITRPEWAKFPSTPQEP-D 433
QY 417 PSPTPTT-TPSP-TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 469
Db 434 DTPYIPTHAPKPSRKTRPKPT--TTTVAATTPVAT-----TTTTHHHH 478
QY 470 HTSETP-----KYTTYVDGHLIKCYKEGDIPIPTNLIHKYLCVFNWGGWVH---- 516
Db 479 HESEKPSQDNQVGSQDTATD---VDCSQEDYLPH-EDCNKYRC-----VHGEAV 526
QY 517 IMPCPGTIW 526
Db 527 LFTCREGTIV 536

RESULT 13
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION:  A human chitinase, its recombinant
; TITLE OF INVENTION:  production, its use for decomposing chitin, its use
; TITLE OF INVENTION:  in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES:  16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Hoffmann & Baron
; STREET:  350 Jericho Turnpike
; CITY:  Jericho

STATE:  New York
COUNTRY:  United States of America
ZIP:  11758
COMPUTER READABLE FORM:
MEDIUM TYPE:  Floppy disk
COMPUTER:  IBM PC compatible
OPERATING SYSTEM:  PC-DOS/MS-DOS
SOFTWARE:  PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER:  US/08/486,839
FILING DATE:  07 - June - 1995
ATTORNEY/AGENT INFORMATION:
NAME:  Baron, Ronald J.
REGISTRATION NUMBER:  29,281
REFERENCE/DOCKET NUMBER:  294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE:  (516) 822-3550
TELEFAX:  (516) 822-3582
INFORMATION FOR SEQ ID NO:  4:
SEQUENCE CHARACTERISTICS:
LENGTH:  466 amino acids
TYPE:  amino acid
STRANDEDNESS:  unknown
TOPOLOGY:  unknown
MOLECULE TYPE:  protein
HYPOTHETICAL:  NO
US-08-486-839-4

Query Match      24.5%; Score 738; DB 2; Length 466;
Best Local Similarity 34.2%; Pred. No. 6.6e-51;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;

QY 15 RIVCVGWSVYHKVDP-YTTIEDDPFKCTHMYGFAKIDKYKTYIQVDFPQYDNNHNSW 73
Db 23 KLVCYFTWQVYRQGEARFLPKDLPSLCTHLYAFAGMTNHQLST-----TEW 71
QY 74 -BKRGYERFNNLRKNPELTWISLGGWYEGSEKYSDMAANPTYRQFIQSVDLFLQYK 132
Db 72 NDETLYQEFNGLKKNPKLKTLLAIGGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYS 131
QY 133 FDGLDLDWYPSGR-LGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAAVSPGDK 185
Db 132 FDGLDLDWYPSGR-LGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAAVSPGDK 190
QY 186 IDRAYDIKELNKLFDWNNVMTYDVGWENFYGHNAFLYKRPDETDELHTYFNVTMHY 245
Db 191 VDAGYEVNKLQNLDFVNLMAVDYFHGSWEKVTGHSNPLYKQESGAAS-LNVDAAVQQ 249
QY 246 YLANGATRDKLVMGVFPYGRAMWEDRSK-LKLQ---DPAKGMSPGPGFISGEEGLSYIE 305
Db 250 WLQKGTTPASKLILGMPYGRSFTLASSSDTRVGPATGSGTGPPTKEGGLAYVEVGSW 309
QY 306 FQKEEWHIOYDEYNNAPYGYNDKTIWGYDDLASISCKLAFKELGVSGVMVWSLENDPK 365
Db 310 KGATKRIQ-DQ--KVPIFRDNQWGVDFDVESEFKTSYKQKGLGAMWALDLDFA 366
QY 366 GHCGPKNPLLNKVNHNMGDEKNSFCILGPSTTTTPPTTPTTPTTPTTPTTPTTPTT 425
Db 367 G-----FSCNQ-----RYPLIQTLRQELSLVLPSTG 394
QY 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450
Db 395 PELEVP-KEQQSEPEHGFSPGQDT 418

RESULT 14
US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION:  A Human Chitinase, Its Recombinant
```

TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morriss, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-151-011-4

Query Match 24.5%; Score 738; DB 3; Length 466;
Best Local Similarity 34.2%; Pred. No. 6.6e-51;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;
QY 15 RIVCVYGVTSVYHKVDP-YTIEDIDPFKCTHLMYGFPAKIDYKTIQVDFDYQDNNNSW 73
DB 23 KLVCYFTTNAQYRQGEARFLPKDLSLCTHLLIYAFAGMTHQLST-----TEW 71
QY 74 -EKRGYERFNNRLKLNKPELTMTISLGGWYEGSEKYSDMAANPTVROQFIQSVDLFLQYK 132
DB 72 NDETLIQEFNGLKKNNPKLTKLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYS 131
QY 133 FDGLDLDWEYPGSRIGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAAVSPGDK 185
DB 132 FDGLDLDWEYPGSQ-GSPAVDKERFTTLVQDLANAFQQAQTSQGERLLLSAAVPAQGY 190
QY 186 IDRAYDIKELNKLFWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVTMTHY 245
DB 191 VDAGYEVNDKIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKREESGAAS-LNVDAAVQ 249
QY 246 YLNNAGATDKLVNMGVFFYGRAMSIEDRSKLGLDPAKMSPPGFTSGEGVLSYIELCOL 305
DB 250 WLQKGTTPASKLILGMPTTYGRSTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCW 309
QY 306 FQKEWHIQYDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVWVWVSLNDDFK 365
DB 310 KGATKQRIQ-DQ--KVPYIFRDNQWVGFDVDFESFKTKVSYLKQKGLGAMWALDLDFA 366
QY 366 GHCGPKNPLNKNVNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPTTPTTPTT 425
DB 367 G-----FSCNQ-----RYPLIQTLRQELSLPVLPSGT 394
QY 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450
DB 395 PELEVP-KFGQSEPEHGSPGQDT 418

RESULT 15
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-039-198A-2

Query Match 24.5%; Score 738; DB 3; Length 466;
Best Local Similarity 34.2%; Pred. No. 6.6e-51;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;
QY 15 RIVCVYGVTSVYHKVDP-YTIEDIDPFKCTHLMYGFPAKIDYKTIQVDFDYQDNNNSW 73
DB 23 KLVCYFTTNAQYRQGEARFLPKDLSLCTHLLIYAFAGMTHQLST-----TEW 71
QY 74 -EKRGYERFNNRLKLNKPELTMTISLGGWYEGSEKYSDMAANPTVROQFIQSVDLFLQYK 132
DB 72 NDETLIQEFNGLKKNNPKLTKLLAIGGWNFGTKFTDMVATANNRQTFVNSAIRFLKYS 131
QY 133 FDGLDLDWEYPGSRIGNPKIDKQNYLALVRELKDAFEPHG-----YLLTAAVSPGDK 185
DB 132 FDGLDLDWEYPGSQ-GSPAVDKERFTTLVQDLANAFQQAQTSQGERLLLSAAVPAQGY 190
QY 186 IDRAYDIKELNKLFWNNVMTYDHGGWENFYGHNAPLYKRPDETDELHTYFNVTMTHY 245
DB 191 VDAGYEVNDKIAQNLDVFNLMAYDFHGSWEKVTGHNSPLYKREESGAAS-LNVDAAVQ 249
QY 246 YLNNAGATDKLVNMGVFFYGRAMSIEDRSKLGLDPAKMSPPGFTSGEGVLSYIELCOL 305
DB 250 WLQKGTTPASKLILGMPTTYGRSTLASSSDTRVGAPATGSGTGPFTKEGMLAYEVCW 309
QY 306 FQKEWHIQYDEYNNAPYGVNDKIWGYDDLASISCKLAFKELGVSGVWVWVSLNDDFK 365
DB 310 KGATKQRIQ-DQ--KVPYIFRDNQWVGFDVDFESFKTKVSYLKQKGLGAMWALDLDFA 366
QY 366 GHCGPKNPLNKNVNMINGDEKNSFECILGSPSTTTTPTTPTTPTTPTTPTTPTTPTT 425
DB 367 G-----FSCNQ-----RYPLIQTLRQELSLPVLPSGT 394
QY 426 PSPTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 450

Db 395 PELEVP-KFGQSPSEPHGSPGQDT 418

Search completed: March 22, 2004, 07:04:01
Job time : 44.2018 secs

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

RESULT 1

```
Query Match      100.0%; Score 3107; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

| | | | | | | | | | | | | | | | | |
|-----|-------|-----|---------|---------|------|-----|-----|------|-----|-----|-----|------|-----|-----|-----|-----|
| 1 | MKTYI | ILS | IMACIGL | WNASIKR | DNDY | SKN | PMR | IVCV | YGT | WSV | YHK | VD | PTY | IED | DFK | 60 |
| 1 | MKTYI | ILS | IMACIGL | WNASIKR | DNDY | SKN | PMR | IVCV | YGT | WSV | YHK | VD | PTY | IED | DFK | 60 |
| 61 | CTHLM | YGF | AKID | EYK | YTI | QV | FD | YQ | DDN | NS | WEK | GYER | FNN | LR | LKN | PEL |
| 61 | CTHLM | YGF | AKID | EYK | YTI | QV | FD | YQ | DDN | NS | WEK | GYER | FNN | LR | LKN | PEL |
| 121 | EGSEK | YSD | MA | NP | TY | QQ | FI | Q | S | V | L | D | F | Q | EYK | FD |
| 121 | EGSEK | YSD | MA | NP | TY | QQ | FI | Q | S | V | L | D | F | Q | EYK | FD |
| 181 | REL | KD | A | EP | H | G | YLL | T | AA | V | S | PK | D | I | R | A |
| 181 | REL | KD | A | EP | H | G | YLL | T | AA | V | S | PK | D | I | R | A |
| 241 | PLY | KR | P | E | T | E | L | E | T | F | N | V | N | T | M | H |
| 241 | PLY | KR | P | E | T | E | L | E | T | F | N | V | N | T | M | H |

SUMMARIES

| Result No. | Query | | | ID | Description |
|------------|-------|-------|--------|----|-------------------|
| | Score | Match | Length | | |
| 1 | 3107 | 100.0 | 555 | 4 | US-09-292-2235-15 |
| 2 | 3107 | 100.0 | 555 | 4 | US-09-292-2235-18 |
| 3 | 3014 | 97.0 | 536 | 4 | US-09-292-2235-21 |
| 4 | 2542 | 81.8 | 509 | 4 | US-09-292-2235-35 |
| 5 | 2542 | 81.8 | 509 | 4 | US-09-292-2235-38 |
| 6 | 2475 | 79.7 | 490 | 4 | US-09-292-2235-41 |
| 7 | 838.5 | 27.0 | 554 | 2 | US-08-524-051-2 |
| 8 | 838.5 | 27.0 | 554 | 3 | US-09-052-778-16 |
| 9 | 815.5 | 26.2 | 583 | 4 | US-09-545-814-2 |
| 10 | 815.5 | 26.2 | 583 | 4 | US-09-545-814-5 |
| 11 | 812.5 | 26.2 | 635 | 4 | US-09-545-814-32 |
| 12 | 811.5 | 26.1 | 559 | 4 | US-09-545-814-14 |
| 13 | 738 | 23.8 | 466 | 2 | US-08-486-833-4 |
| 14 | 738 | 23.8 | 466 | 3 | US-09-151-011-4 |
| 15 | 738 | 23.8 | 466 | 3 | US-09-039-198A-2 |
| 16 | 738 | 23.8 | 466 | 4 | US-09-343-623-4 |
| 17 | 738 | 23.8 | 466 | 4 | US-08-877-599-2 |
| 18 | 738 | 23.8 | 466 | 4 | US-09-267-574-2 |
| 19 | 732.5 | 23.6 | 387 | 2 | US-08-486-833-6 |
| 20 | 732.5 | 23.6 | 387 | 3 | US-09-151-011-6 |
| 21 | 732.5 | 23.6 | 387 | 4 | US-09-343-623-6 |
| 22 | 732 | 23.6 | 466 | 3 | US-09-039-198A-4 |
| 23 | 732 | 23.6 | 466 | 4 | US-08-877-599-4 |
| 24 | 732 | 23.6 | 466 | 4 | US-09-267-574-4 |
| 25 | 721.5 | 23.2 | 373 | 3 | US-09-039-198A-14 |
| 26 | 721.5 | 23.2 | 373 | 3 | US-09-039-198A-15 |
| 27 | 721.5 | 23.2 | 373 | 4 | US-08-877-599-14 |

Db 241 PLYKRPDETDLHTYFNVNTHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPA 300
QY 301 KMSPPGFIISGEGVLSYIELCOLFOKEEWHIYQDEYNNAPYNDKIWGVYDDLASISC 360
Db 301 KMSPPGFIISGEGVLSYIELCOLFOKEEWHIYQDEYNNAPYNDKIWGVYDDLASISC 360
QY 361 KLAFLKELGSGVMVWSLENDDFKGCHGCPKNPLLNKVHNMINGDEKNSFECILGSPSTTP 420
Db 361 KLAFLKELGSGVMVWSLENDDFKGCHGCPKNPLLNKVHNMINGDEKNSFECILGSPSTTP 420
QY 421 TPPTTPT 480
Db 421 TPPTTPT 480
QY 481 TPSPTTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGVWVHIMPCP 540
Db 481 TPSPTTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGVWVHIMPCP 540
QY 541 PGTIWCQEKLTICIG 555
Db 541 PGTIWCQEKLTICIG 555

RESULT 2

US-09-292-225-18
; Sequence 18, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-18

Query Match 100.0%; Score 3107; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKTIYALISIMACIGLNNASIKRDNDYSKNPRLVYCVGTWSVYHKVDPTTIEDIDPFK 60
Db 1 MKTIYALISIMACIGLNNASIKRDNDYSKNPRLVYCVGTWSVYHKVDPTTIEDIDPFK 60
QY 61 CTHLMYGFADIKYKTIQVDFPDYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGGWY 120
Db 61 CTHLMYGFADIKYKTIQVDFPDYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGGWY 120
QY 121 EGSEKYSMDMAANPTYRQOFIQSVLDFLOEYKFDGLDWEYPSRGLGNPKIDKQNYALV 180
Db 121 EGSEKYSMDMAANPTYRQOFIQSVLDFLOEYKFDGLDWEYPSRGLGNPKIDKQNYALV 180
QY 181 RELKDAPEPFGYLLTAASVSGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240
Db 181 RELKDAPEPFGYLLTAASVSGKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNA 240

QY 241 PLYKRPDETDLHTYFNVNTHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPA 300
Db 241 PLYKRPDETDLHTYFNVNTHYLLNNGATRDKLVGVFPFYGRAWSIEDRSKILGDPA 300
QY 301 KMSPPGFIISGEGVLSYIELCOLFOKEEWHIYQDEYNNAPYNDKIWGVYDDLASISC 360
Db 301 KMSPPGFIISGEGVLSYIELCOLFOKEEWHIYQDEYNNAPYNDKIWGVYDDLASISC 360
QY 361 KLAFLKELGSGVMVWSLENDDFKGCHGCPKNPLLNKVHNMINGDEKNSFECILGSPSTTP 420
Db 361 KLAFLKELGSGVMVWSLENDDFKGCHGCPKNPLLNKVHNMINGDEKNSFECILGSPSTTP 420
QY 421 TPPTTPT 480
Db 421 TPPTTPT 480
QY 481 TPSPTTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGVWVHIMPCP 540
Db 481 TPSPTTTEHTSETPKYTYVDGHLIKCYKEGDIHPHTNIHKYLVCFVNGVWVHIMPCP 540
QY 541 PGTIWCQEKLTICIG 555
Db 541 PGTIWCQEKLTICIG 555

RESULT 3

US-09-292-225-21
; Sequence 21, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-21

Query Match 97.0%; Score 3014; DB 4; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.8e-233;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 SIKRDNDYSKNPRLVYCVGTWSVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 79
Db 1 SIKRDNDYSKNPRLVYCVGTWSVYHKVDPTTIEDIDPFKCTHLMYGFADIKYKTIQ 60
QY 80 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGGWYEGSEKYSMDMAANPTYRQOF 139
Db 61 VFDPYQDDNHSWEKRGYERFNNLRLKNPELTMTISLGGWYEGSEKYSMDMAANPTYRQOF 120
QY 140 IQSVLDFLOEYKFDGLDWEYPSRGLGNPKIDKQNYALVRELKDAPEPFGYLLTAASV 199
Db 121 IQSVLDFLOEYKFDGLDWEYPSRGLGNPKIDKQNYALVRELKDAPEPFGYLLTAASV 180
QY 200 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNVN 259
Db 181 PGKDKIDRAYDIKELNKLFDWNNVMTYDYGHWENFYGHNAPLYKRPDETDLHTYFNVN 240

Db 241 PLYKREDEDELHTYFNVNYTHMYLNNATRDKLVNGVFFYGRAWSIEDRSKVLGDP 300
QY 301 KGMSPGFTSGEGLVSYIELCOLFOKEBWHIOYDEYNNAPYGYNDKIWVGDDDLASISC 360
Db 301 KGMSPGFTSGEGLVSYIELCOLFOKEBWHIOYDEYNNAPYGYNDKIWVGDDDLASISC 360
QY 361 KLAFLKELGVSGVWVWLSLNDPKHCGKPNPLNKHVNMINGDEKNSFECILGSPSTTP 420
Db 361 KLAFLKELGVSGVWVWLSLNDPKHCGKPNPLNKHVNMINGDEKNSFECILGSPSTTP 420
QY 421 TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
Db 421 TPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 480
QY 481 TPSPPTTHTSETPKYTYVVDGHLIKCYKEGDIPHTNIHKYLVCFFV---NGGWWVHIM 537
Db 435 ---PTTDTSTSETPKYTYVVDGHLIKCYKEGDIPHTNIHKYLVCFFV---NGGWWVHIM 491
QY 538 PDPGFTWCEKLTCTIGE 555
Db 492 DCPKGTWHTATLKNICQE 509

RESULT 6
US-09-292-225-41
; Sequence 41, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley W.
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-41

Query Match 79.7%; Score 2475; DB 4; Length 490;
Best Local Similarity 82.5%; Pred. No. 2.5e-190;
Matches 443; Conservative 21; Mismatches 21; Indels 52; Gaps 2;
QY 22 KRDNDYKSNMRYICVGTWVSHKVDYPTIEDIDPKCTHLMYGFADIKYKTIQVF 81
Db 3 KRDHNNYSKNMRYICVGTWVSHKVDYPTIEDIDPKCTHLMYGFADIKYKTIQVF 62
QY 82 DPYQDNHNSWEKRGVERENNLKKNPELTMTISLGGWYEGSEKYSDMAANPTTRQIQ 141
Db 63 DPFQDDHNSWEKRGVERENNLKKNPELTMTISLGGWYEGSEKYSDMAANPTTRQIQ 122
QY 142 SVLDFLOEYKFGDLDDWEPGSRGNPKIDKQNYLALVRELKDAFEHGYLLTAASPG 201
Db 123 SVLDFLOEYKFGDLDDWEPGSRGNPKIDKQNYLALVRELKDAFEHGYLLTAASPG 182
QY 202 KDKIDPRAYDIKELNKLFDWMVNTYDYGWENFYNCHNAPLYKRPDETDDELHTYFNVNYT 261
Db 183 KDKIDVAYELKELNQLFDWMVNTYDYGWENFYNCHNAPLYKRPDETDDELHTYFNVNYT 242

QY 262 MHYTLNNGATRDKLVNGVPPYGRAWSIEDRSKVLGDPKMGSPPGFISGEGVLSYIEL 321
Db 243 MHYTLNNGATRDKLVNGVPPYGRAWSIEDRSKVLGDPKMGSPPGFISGEGVLSYIEL 302
QY 322 COLFOKEBWHIOYDEYNNAPYGYNDKIWVGDDDLASISCKLAFKELGVSGVWVWLSLND 381
Db 303 COLFOKEBWHIOYDEYNNAPYGYNDKIWVGDDDLASISCKLAFKELGVSGVWVWLSLND 362
QY 382 DFKHCGKPNPLNKHVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPTTPTT 441
Db 363 DFKHCGKPNPLNKHVNMINGDEKNSFECILGSPSTTPTTPTTPTTPTTPTTPTTPTTPTT 415
QY 442 TTPTSPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 501
Db 416 ---PTTDTSTSETPKYTYVVDGHLIKCYKEGDIPHTNIHKYLVCFFV---NGGWWVHIM 433
QY 502 GHLIKCYKEGDIPHTNIHKYLVCFFV---NGGWWVHIMCPDGTWCEKLTCTIGE 555
Db 434 GHLIKCYKQGLPHTDVKYLVCEYATPENGWVHIMDCPKGTWHTATLKNICQE 490

RESULT 7
US-08-524-051-2
; Sequence 2, Application US/08524051
; Patent No. 5866788
; GENERAL INFORMATION:
; APPLICANT: Kramer, Karl J.
; APPLICANT: Muthukrishnan, Subbaratnam
; APPLICANT: Choi, Hee Kyung
; APPLICANT: Corpuz, Lolita
; APPLICANT: Gopalakrishnan, Bhuvana
; TITLE OF INVENTION: RECOMBINANT CHITINASE AND USE THEREOF AS
; TITLE OF INVENTION: A BIOCIDES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Blvd., Suite 400
; CITY: Kansas City
; STATE: MO
; COUNTRY: USA
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/524,051
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Collins, John M.
; REGISTRATION NUMBER: 26,262
; REFERENCE/DOCKET NUMBER: 22875-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (816)474-9050
; TELEFAX: (816)474-9057
; TELEX: 434-363
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-524-051-2

Query Match 27.0%; Score 838.5; DB 2; Length 554;
Best Local Similarity 35.4%; Pred. No. 4.8e-59;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;
QY 6 ALLISIMACIGLMMASIKRDNDYKSNMRYICVGTWVSHK-VDPYTTIEDIDPKCTH 64
Db 3 ATLATLAVLATAV-----QSDSRARIVCYFNWAVRPGVGRYGIEDIPVEKCTHI 55

```

QY 65 MYGFAKIDEXYKTIQVEDPYQDDNHNHSEKRGYERFNNRLKNPELTMTTISLGGWEGSE 124
Db 56 IYSGFVTEGSEVLIDFELD-----VDKNGFRNFTSLRSSHPYSVKFWAVGGWAGSS 110
QY 125 KYSDMAANPTYROQFTQSVLDFLOEYKFDGLDLDWEYPGSR-LGNPKIDKQNYLALVREL 183
Db 111 KYSHMVAQKSTRMSFIRSVVFLKDYDFDGLDLDWEYPGAADRGGSFSDKDFLYLVQEL 170
QY 184 KDAP--EPHGILLTAASVSGKIDRAYDIKELNKLFDMMNMTYDHYHGWENFYGHAP 241
Db 171 RRAIRVKGWELTAAPLANFRLMEGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDETDELHTYFNNVYNNNGATRDKLVMGVPPFYGRWSIED-----290
Db 231 LYKRPD-QWAEKLNVDGLHWEKGCPSNKLVGIPFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLKGDPAGKMPGPFISGEEGVLSYIELCOLFOKEE--WHIQDYEYNNAPYGYNDKI 348
Db 290 NKEAGGDPAPYNTATG-----WAYYEICTEVDKDDSGWTKKWEQKCPYAYKGTQ 342
QY 349 WVGVDLLASISCKLAFELKELGVSVMWVWSLENDDFKHCGRKPNLLKNVHNMINGDERNS 408
Db 231 LYKRPD-QWAEKLNVDGLHWEKGCPSNKLVGIPFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLKGDPAGKMPGPFISGEEGVLSYIELCOLFOKEE--WHIQDYEYNNAPYGYNDKI 348
Db 290 NKEAGGDPAPYNTATG-----WAYYEICTEVDKDDSGWTKKWEQKCPYAYKGTQ 342
QY 409 FECILGSPSTTPTPTPTPT-----TPTPTPS-----PTPTPTPTPTPTPTPTSP 456
Db 397 -----SYTVPPHTENTTPTPEWARPPSTPSDPSGDPPTTTTAKPASTTKTKTKTT 449
QY 457 TPTPTPTPTPT-----PTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 505
Db 450 TTTTAKPPQSVIDBENDINVRPEKPEQPEVEVP-PTENE-----VDGSEI 497
QY 506 KYKEGD-IPHTNIHKYLVCVFNGVGGWVHIMPCPPGTIWCQELTC 552
Db 498 -CNSDDQYIPDKKCHDKYRWC--VNGE--AMQFSCQHGTVFNVELNVC 540

```

RESULT 8

```

US-09-052-778-16
; Sequence 16, Application US/09052778A
; Patent No. 6060590
; GENERAL INFORMATION:
; APPLICANT: Bryant, Peter J.
; APPLICANT: Kawamura, Kazuo
; TITLE OF INVENTION: CHITINASE RELATED PROTEINS AND METHODS
; FILE REFERENCE: OF USE
; CURRENT APPLICATION NUMBER: US/09/052, 778A
; CURRENT FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Manduca sexta
US-09-052-778-16

```

```

Query Match 27.0%; Score 838.5; DB 3; Length 554;
Best Local Similarity 35.4%; Pred. No. 4.8e-59;
Matches 208; Conservative 79; Mismatches 210; Indels 91; Gaps 20;

QY 6 AILSMACIGLMASTIKRDHNDYKNPMRIVCVGTWWSVYHK-VDPYTTIEDIDPFKCTHL 64
Db 3 ATUATLAVLATAV-----QDSRARIVCVFNSWAVRPGVGRYGIEDIPVEKCTHI 55
QY 65 MYGFAKIDEXYKTIQVEDPYQDDNHNHSEKRGYERFNNRLKNPELTMTTISLGGWEGSE 124
Db 56 IYSGFVTEGSEVLIDFELD-----VDKNGFRNFTSLRSSHPYSVKFWAVGGWAGSS 110
QY 125 KYSDMAANPTYROQFTQSVLDFLOEYKFDGLDLDWEYPGSR-LGNPKIDKQNYLALVREL 183
Db 111 KYSHMVAQKSTRMSFIRSVVFLKDYDFDGLDLDWEYPGAADRGGSFSDKDFLYLVQEL 170

```

```

QY 184 KDAP--BPHGYLLTAASVSGKIDRAYDIKELNKLFDMMNMTYDHYHGWENFYGHAP 241
Db 171 RRAIRVKGWELTAAPLANFRLMEGYHVPCLQELDAIHVMSYDLRGNWAGFADVHSP 230
QY 242 LYKRPDETDELHTYFNNVYNNNGATRDKLVMGVPPFYGRWSIED-----290
Db 231 LYKRPD-QWAEKLNVDGLHWEKGCPSNKLVGIPFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLKGDPAGKMPGPFISGEEGVLSYIELCOLFOKEE--WHIQDYEYNNAPYGYNDKI 348
Db 290 NKEAGGDPAPYNTATG-----WAYYEICTEVDKDDSGWTKKWEQKCPYAYKGTQ 342
QY 349 WVGVDLLASISCKLAFELKELGVSVMWVWSLENDDFKHCGRKPNLLKNVHNMINGDERNS 408
Db 231 LYKRPD-QWAEKLNVDGLHWEKGCPSNKLVGIPFYGRSFTLSAGNNYGLGTFI 289
QY 291 RSKLKGDPAGKMPGPFISGEEGVLSYIELCOLFOKEE--WHIQDYEYNNAPYGYNDKI 348
Db 290 NKEAGGDPAPYNTATG-----WAYYEICTEVDKDDSGWTKKWEQKCPYAYKGTQ 342
QY 409 FECILGSPSTTPTPTPTPT-----TPTPTPS-----PTPTPTPTPTPTPTPTSP 456
Db 397 -----SYTVPPHTENTTPTPEWARPPSTPSDPSGDPPTTTTAKPASTTKTKTKTT 449
QY 457 TPTPTPTPTPT-----PTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 505
Db 450 TTTTAKPPQSVIDBENDINVRPEKPEQPEVEVP-PTENE-----VDGSEI 497
QY 506 KYKEGD-IPHTNIHKYLVCVFNGVGGWVHIMPCPPGTIWCQELTC 552
Db 498 -CNSDDQYIPDKKCHDKYRWC--VNGE--AMQFSCQHGTVFNVELNVC 540

```

RESULT 9

```

US-09-545-814-2
; Sequence 2, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-2

```

```

Query Match 26.2%; Score 815.5; DB 4; Length 583;
Best Local Similarity 34.6%; Pred. No. 3.6e-57;
Matches 198; Conservative 94; Mismatches 205; Indels 75; Gaps 23;

QY 7 ILSIMACIGLMASTIKRDHNDYKNPMRIVCVGTWWSVYHK-VDPYTTIEDIDPFKCTHL 65
Db 6 LLAVLCAIAISSINTV-----EASQKARIVCVFNSWAVRPGVGRYGIEDIPVLCCTHIV 61
QY 66 YGFAKIDEXYKTIQVEDPYQDDNHNHSEKRGYERFNNRLKNPELTMTTISLGGWEGSEK 125
Db 62 YSFTGVDDKDSVSLVIDPELDIDN-----GFKNFTLRLKTHPNVQLQIAGVGAEGGKK 116
QY 126 YSDMAANPTYROQFTQSVLDFLOEYKFDGLDLDWEYPGSR-LGNPKIDKQNYLALVREL 184
Db 117 YSTWVAEKRSASFIRSVVDFMNEYKFDGLDLDWEYPGAADRGGSFSDKDFLYFVQELR 176
QY 185 DAFEPHG--YLLTAASVSGKIDRAYDIKELNKLFDMMNMTYDHYHGWENFYGHAP 242
Db 177 RAFPKNQKNWEITWAVPIAKFRLOEGVHVPCLQELDAIHVMSYDLRGNWAGFADHSP 236
QY 243 YKRPDETDELHTYFNNVYNNNGATRDKLVMGVPPFYGRWSIEDRSK-LKLG---D 298

```

Db 237 YRRPHD-QYAYEKLNVNDGLQLVDMGCPANKLVGVFPYGRSFTLSNSKDYRLGTYN 295
Qy 299 PAKGMSPPGFSIGBEGVLSYIELCOLFQ--KEEMHIQYDEYNAPGYNDKIWGVYDDLA 356
Db 296 KEAGGEGPGYTNATGFSYVEICLEVDPSKGTWKWDHKGKVPYAYKGNQWGVYEDPK 355
Qy 357 SISCKLAFKELGVSGVMVWSLENDDPKHCG-PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 356 SVALKMEFIKSKGYGAMTWAIDMDDFQGVCSDDKHTLAVIMHDYM---KNYIVPEFDS 411
Qy 416 STTTPTP--TTTPTPTPTPTPTPTPTT--TPSPT--TPPTTPTPTPTPTPTPTPTPT 471
Db 412 SRIIPREWAKPPSPSQEP-DDTPIPTTHAPKPSRKPKPT--TTTVAATTPVAT 468
Qy 472 TPTAPTSTPSPTTTE-----HTSETP-----KYTTVDGHLIKCYKEGDIPHPT 517
Db 469 -----TTTEHHHHHEEBEKPSEQDNQVGSQDTATD---VDCSQEDYLPH-E 511
Qy 518 NIHXYLVCBFVNGWVWH-----IMPCPPGTIW 545
Db 512 DCKNYRC-----VHGEAVLFTCREGTIV 535

RESULT 10
US-09-545-814-5
; Sequence 5, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-5

Query Match 26.2%; Score 815.5; DB 4; Length 583;
Best Local Similarity 34.6%; Pred. No. 3.6e-57;
Matches 198; Conservative 94; Mismatches 205; Indels 75; Gaps 23;

Qy 7 ILSIMACIGLMNASIKRDHNDYSKNPMRIVCVVGTWSVYHK-VDPYTTIEDIDPFKCTHLM 65
Db 6 LLAVLCAIAISSINTV-----EASDQKARIVCVFSNVAVRPGIRYGIEDIPVDLCTHIV 61
Qy 66 YGFAKIDYKYTIQVDPYQDDNHNHNSWEKRGVERFNNLRKNPELTTMISLGGWYEGSEK 125
Db 62 YSFIGVDDKDSVSLVIDPELDIDDN-----GFKFTNLRKHHPNVKLQIAGVGAEGGKK 116
Qy 126 YSDMAANTYQQFQSVLDFLOEYKFDGLDWEYPGSR-LGNPKIDKQNYLALVRELK 184
Db 117 YSTMVAEKRSKSAFIRSVVDFMNEYKFDGFDLWEYPGAADRGGSFSDKDKFLYFVQELR 176
Qy 185 DAFEPHG--YLLTAASVPGKIDRAYDIKELNKLFDMMVNTYDYGWENFYGHNAFL 242
Db 177 RAFNKGQKWEITWAVPIAKFRLQEGYHVPCLCELLDAIHVMSYDLRGNWAGFADTHSPL 236
Qy 243 YKRPDDELHYFNVNTYTHYLNAGTRDKLVGVFPYGRWSIEDRSK-LKLG---D 298
Db 237 YRRPHD-QYAYEKLNVNDGLQLVDMGCPANKLVGVFPYGRSFTLSNSKDYRLGTYN 295
Qy 299 PAKGMSPPGFSIGBEGVLSYIELCOLFQ--KEEMHIQYDEYNAPGYNDKIWGVYDDLA 356
Db 296 KEAGGEGPGYTNATGFSYVEICLEVDPSKGTWKWDHKGKVPYAYKGNQWGVYEDPK 355
Qy 357 SISCKLAFKELGVSGVMVWSLENDDPKHCG-PKNPLLNKVNMMINGDEKNSFECILGP 415

Db 356 SVALKMEFIKSKGYGAMTWAIDMDDFQGVCSDDKHTLAVIMHDYM---KNYIVPEFDS 411
Qy 416 STTTPTP--TTTPTPTPTPTPTPTPTT--TPSPT--TPPTTPTPTPTPTPTPTPTPT 471
Db 412 SRIIPREWAKPPSPSQEP-DDTPIPTTHAPKPSRKPKPT--TTTVAATTPVAT 468
Qy 472 TPTAPTSTPSPTTTE-----HTSETP-----KYTTVDGHLIKCYKEGDIPHPT 517
Db 469 -----TTTEHHHHHEEBEKPSEQDNQVGSQDTATD---VDCSQEDYLPH-E 511
Qy 518 NIHXYLVCBFVNGWVWH-----IMPCPPGTIW 545
Db 512 DCKNYRC-----VHGEAVLFTCREGTIV 535

RESULT 11
US-09-545-814-32
; Sequence 32, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE REFERENCE: FC-5-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 32
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Tagged Ctenocephalides felis
US-09-545-814-32

Query Match 26.2%; Score 812.5; DB 4; Length 635;
Best Local Similarity 34.6%; Pred. No. 7e-57;
Matches 198; Conservative 93; Mismatches 206; Indels 75; Gaps 23;

Qy 7 ILSIMACIGLMNASIKRDHNDYSKNPMRIVCVVGTWSVYHK-VDPYTTIEDIDPFKCTHLM 65
Db 7 LLAVLCAIAISSINTV-----EASDQKARIVCVFSNVAVRPGIRYGIEDIPVDLCTHIV 62
Qy 66 YGFAKIDYKYTIQVDPYQDDNHNHNSWEKRGVERFNNLRKNPELTTMISLGGWYEGSEK 125
Db 63 YSFIGVDDKDSVSLVIDPELDIDDN-----GFKFTNLRKHHPNVKLQIAGVGAEGGKK 117
Qy 126 YSDMAANTYQQFQSVLDFLOEYKFDGLDWEYPGSR-LGNPKIDKQNYLALVRELK 184
Db 118 YSTMVAEKRSKSAFIRSVVDFMNEYKFDGFDLWEYPGAADRGGSFSDKDKFLYFVQELR 177
Qy 185 DAFEPHG--YLLTAASVPGKIDRAYDIKELNKLFDMMVNTYDYGWENFYGHNAFL 242
Db 178 RAFNKGQKWEITWAVPIAKFRLQEGYHVPCLCELLDAIHVMSYDLRGNWAGFADTHSPL 237
Qy 243 YKRPDDELHYFNVNTYTHYLNAGTRDKLVGVFPYGRWSIEDRSK-LKLG---D 298
Db 238 YRRPHD-QYAYEKLNVNDGLQLVDMGCPANKLVGVFPYGRFPTLSNSKDYRLGTYN 296
Qy 299 PAKGMSPPGFSIGBEGVLSYIELCOLFQ--KEEMHIQYDEYNAPGYNDKIWGVYDDLA 356
Db 297 KEAGGEGPGYTNATGFSYVEICLEVDPSKGTWKWDHKGKVPYAYKGNQWGVYEDPK 356
Qy 357 SISCKLAFKELGVSGVMVWSLENDDPKHCG-PKNPLLNKVNMMINGDEKNSFECILGP 415
Db 357 SVALKMEFIKSKGYGAMTWAIDMDDFQGVCSDDKHTLAVIMHDYM---KNYIVPEFDS 412
Qy 416 STTTPTP--TTTPTPTPTPTPTPTPTT--TPSPT--TPPTTPTPTPTPTPTPTPTPT 471
Db 413 SRIIPREWAKPPSPSQEP-DDTPIPTTHAPKPSRKPKPT--TTTVAATTPVAT 469

Qy 472 TPTAPTSTPSPTTTE-----HTSETP-----KYTTYVDGHLIKYKEDGIPHPT 517
Db 470 -----TTTEHHHHEEKEKPSQDNQVGSQDTATD---VDCSQEDYLFH-E 512
Qy 518 NIHKYLVCFVNGGWWH-----IMPCPPGTIW 545
Db 513 DCKKYRC-----VHGEAVLFTCREGTIVY 536
RESULT 12
US-09-545-814-14
; Sequence 14, Application US/09545814
; Patent No. 6416977
; GENERAL INFORMATION:
; APPLICANT: Becher, Anna M.
; TITLE OF INVENTION: FLEA CHITINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-S-C1
; CURRENT APPLICATION NUMBER: US/09/545,814
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,833
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Ctenocephalides felis
US-09-545-814-14
Query Match 26.1%; Score 811.5; DB 4; Length 559;
Best Local Similarity 35.8%; Pred. No. 7.1e-57;
Matches 193; Conservative 86; Mismatches 193; Indels 71; Gaps 22;
Qy 34 RIVCVGTWVYTHK-VDPYTIEDIDPFKTHLMYGFAXIDYKTYIQVDFPYQDDNHNWS 92
Db 5 RIVCVFSNWAYRPGIGRYGIEDIPVLCITHIVYSGVDDKDSVLVIDPELDIDDN-- 62
Qy 93 EKGYERFNNRLKNPBLTMTISLGGWYEGSEKYSMDAANPTYRQFIQSVLDFLOEYKF 152
Db 63 ---GFKNFTNLRKIHNPVKLIQIAGVGAEGKYSYTWAEKRKRSAFIRSVDFMNEYKF 119
Qy 153 DGLDLWEYFGSR-LGNPKDKQNYLALVRELKDAFEPHG--YLLTAASVSPGDKIDRAY 209
Db 120 DGFPLDWEYFGAARDGGSFSDCKFLYFVQELRAFNFKNQKWEITMAVFIARFLQEGY 179
Qy 210 DIKELNKLFDWMNVMTYDVGWENFYGHNAPLYKRPDETDLHTYFNVNTMHHYLNNG 269
Db 180 HVPCLCLLDIAHVMYSYDLRGNWAGFADTHSPLYRRPHD-QYAYEKLNVNDQLQWMDWG 238
Qy 270 ATRDKLVMGVFPFGRAWSIEDRSK-LKLG---DPAKMSPPGFGISGEGVLSYIELCOLF 325
Db 239 CPANKLVGVFPFGRSFTLSNSKDYELGTYNKEAGGEGPGTYTNATGFSIYEICLEV 298
Qy 326 Q--KEEMHIQYDEYNAPYNDKINWGYDDLASISCKLAFKLKELGVSGVMVMSLENDPF 383
Db 299 DDPKSGWTKWDEHGKVPYAYKGNQWGYEDPKSVALKMEFIKSKGVGGAMTWAIDMDDF 358
Qy 384 KHCG-KPNFLNKNVHNMINGDEKNSFECILGSTTTPP--TTTTTPTTTTTPSPPT 440
Db 359 QGVCSDDKHTLAVIMHDYM---KNYIVPFDSRITPRPEWAKPSPSTPSQEP-DDTPYI 413
Qy 441 PTT-TPSPPT-TPPTTSPPTTTPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 493
Db 414 PTTTHAPKSPKTRKPKPT--TTTVAATTVPAT-----TTTTHHHHHHEEEK 458
Qy 494 P-----KYTTYVDGHLIKYKEDGIPHTNIHKYLVCFVNGGWWH-----IMPCP 540
Db 459 PSEQDNQVGSQDTATD---VDCSQEDYLFH-EDCKNYRC-----VHGEAVLFTCR 506
Qy 541 PGTIW 545
Db 507 EGTIV 511

RESULT 13
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-486-839-4
Query Match 23.8%; Score 738; DB 2; Length 466;
Best Local Similarity 34.2%; Pred. No. 4.3e-51;
Matches 152; Conservative 85; Mismatches 150; Indels 58; Gaps 11;
Qy 34 RIVCVGTWVYTHKVDY-YTIEDIDPFKTHLMYGFAXIDYKTYIQVDFPYQDDNHNWS 92
Db 23 KLVCFYTNWAQYRQGEARFLPKLDPLSLCTHLIYAFAGTWNHQLST-----TEW 71
Qy 93 -EKGYERFNNRLKNPBLTMTISLGGWYEGSEKYSMDAANPTYRQFIQSVLDFLOEYK 151
Db 72 NDETLYQBFNGLKKNPKLTILLAGTGNWFGTQKFTDMVATANNRQTFVNSAIRFLKYS 131
Qy 152 FDGLDLWEYFGSR-LGNPKDKQNYLALVRELKDAFEPHG-----YLLTAASVSPGDK 204
Db 132 FDGLDLWEYFGSQ-GSPAVDKERFTTLVQDLANAFQQAOTSGKERLLLSAIVPAGQY 190
Qy 205 IDRAYDIKELNKLFDWMNVMTYDVGWENFYGHNAPLYKRPDETDLHTYFNVNTMHHY 264
Db 191 VDAGYEVKIAQLNDFVNLMAVDYFSGWKEVYTGHNPSLYKQESGGAAS-LNVDAVQQ 249
Qy 265 YLNNGATDKLVMGVFPFGRAWSIEDRSKILKLGDPKAGMSPPGFTSGBEGVLSYIELCOL 324
Db 250 WLQKGTASKLILGPTYGRSFTLSSSDTRVGAATGSGTGPPTKEGMLAYEVCWS 309
Qy 325 FQKEWHIYDEYNAPYNDKINWGYDDLASISCKLAFKLKELGVSGVMVMSLENDPF 384
Db 310 KGATKQRIQ-DQ--KVPYIFRDNQWVGDDVBSFRTKYSYLKQKGLGGMVWALDLDFFA 366

| | | | |
|----|-----|--|-----|
| Db | 191 | VDAGYVDKIAQNLDFVNLMAVDFHGSWEKVTHNSPLYKROESGAAS-LNVDAAVQQ | 249 |
| QY | 265 | YLNNGATRDKLYMGVFPYGRAMSIEDRSKILGDPAGMSPPGFYSGEGLVLSYIELCOL | 324 |
| Db | 250 | WLQKGTTPASKLILGMPYGRSFTLASSSDTRVGAPATGSGTGPFTKEGMLAYYEVCSW | 309 |
| QY | 325 | FOKEWHIQDEYNNAPYGYNDKIWVGYYDDLASISCKLAFKLGVSGVMWWSLENDPK | 384 |
| Db | 310 | KGATKQRIQ-DQ--KVEYIFRDNQWYGFDDVESFKTKVSYLKQGLGGAMWALDLDFA | 366 |
| QY | 385 | GHCGPKNPLLNKVNHNINGDEKNSFECILGSPSTTTTPTTTTPTTTTPTTTTPTTTT | 444 |
| Db | 367 | G-----FSCNOG-----RYPLIQTLROELSILPLPSGT | 394 |
| QY | 445 | PSPTTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT | 469 |
| Db | 395 | PELEVP-KPGQPSEPEHGPSPGQDT | 418 |

Search completed: March 22, 2004, 07:04:00
Job time : 46.7332 secs